IBM Smart Analytics Cloud

Understand the business value of a Smart Analytics Cloud

Understand the Smart Analytics Cloud architecture

Implement and manage your own Smart Analytics Cloud

Lydia Parziale
Andrey Avramenko
Simon Chan
Foulques de Valence
Christopher Dziekan
Michael Dziekan

Andrea Greggo
Christian Hagen
Douglas Lin
James Machung
Nicole Roik

Redbooks

ibm.com/redbooks

International Technical Support Organization

IBM Smart Analytics Cloud

September 2010

Note: Before using this information and the product it supports, read the information in "Notices" on page xi.

First Edition (September 2010)

This edition applies to V7.0.0.9 IBM® WebSphere® Application Server Network Deployment (D55WQLL), V8.4.1 IBM Cognos 8 BI (D06ZILL),V9.5 IBM DB2 Enterprise Server (D55J0LL), V6.2.2. ITCAM for Applications (D044LLL), V6.2 FP1 IBM Tivoli Monitoring for Database, V6.2.2 FP1 IBM Tivoli Monitoring Agent for O/S Multiplatform (D5602LL),V7.1 ITCAM Application Diagnostics Agent for HTTP Servers,V7.1 ITCAM Application Diagnostics Agent for WebSphere Applications.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®	InfoSphere™	System z10®
Cognos®	LotusLive™	System z®
DB2 Connect™	OMEGAMON®	Tivoli Enterprise Console®
DB2 Universal Database™	RACF®	Tivoli®
DB2®	Redbooks®	WebSphere®
DirMaint™	Redpaper™	z/Architecture®
FICON®	Redbooks (logo) ®	z/OS®
Global Business Services®	Service Request Manager®	z/VM®
HiperSockets™	System p®	z10™
IBM®	System x®	zSeries®

The following terms are trademarks of other companies:

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

Java, and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Preface

One of the benefits of computing technology is its ability to change the information landscape. The Internet, Wikis, and other technologies make it easy to share ideas rapidly around the world. Our decision making ability, whether it is buying a car for the right price or making a critical business decision, has improved. The challenge, however, is in sorting through all of the information to find what is relevant to the decision you want to make. At the same time, the technology landscape is changing. Transformation to cloud computing is accelerating. We are at a crossroad, one where an explosion in information meets the new cloud computing paradigm. To improve decision making and spark innovation, IBM® offers the Smart Analytics Cloud.

This IBM Redbooks® publication presents a Smart Analytics Cloud. The IBM Smart Analytics Cloud is an IBM offering to enable delivery of business intelligence and analytics at the customer location in a private cloud deployment. The offering leverages a combination of IBM hardware, software and services to offer customers a complete solution that is enabled at their site. In this publication, we provide the background and product information for decision-makers to proceed with a cloud solution.

The content ranges from an introduction to cloud computing to details about our lab implementation. The core of the book discusses the business value, architecture, and functionality of a Smart Analytics Cloud. To provide deeper perspective, documentation is also provided about implementation of one specific Smart Analytics Cloud solution that we created in our lab environment. Additionally, we also describe the IBM Smart Analytics Cloud service offering that can help you create your own Smart Analytics cloud solution that is tailored to your business needs.

The team who wrote this book

This book was produced by a team of specialists from around the world working with the International Technical Support Organization, Poughkeepsie Center.

Lydia Parziale is a Project Leader for ITSO teams in Poughkeepsie, New York. She has domestic and international experience in technology management that includes software development, project leadership, and strategic planning. Her areas of expertise are e-business development and database management technologies. Lydia is a certified Project Management Professional (PMP) and

an IBM Certified IT Specialist with an MBA in Technology Management. She has worked for IBM for 24 years in various technology areas.

Andrey Avramenko is an IBM Level 1 Certified IT Specialist for STG System z Technical Sales in Russia. He has eight years of experience in the IT industry that includes two years at IBM. His areas of expertise are System z hardware, z/VM®, Linux®, and system management software.

Simon Chan is an Advisory IT Architect for IBM Securities Industry Services, FS&I, GTS in Canada. He worked for IBM for 12 years and has more than 15 years of experience in the IT industry prior to IBM. He has a B. Sc. (Hons) degree in Computing Science from the University of Ulster, United Kingdom (UK). Simon is an IBM Certified IT Specialist with proficient technical and IT infrastructure architecture experience in IBM systems and software solutions. His areas of expertise are large-scale multiplatform, multi-clients environment, IBM System z® infrastructure, and WebSphere® technology on System z and distributed systems. He is also a Chartered IT Professional of BCS, The Chartered Institute for IT in the UK, and member of the Institution of Engineering and Technology (IET).

Foulques de Valence is a Web and Security Consultant member of the IBM System z Lab Services team. He provides consulting services in the USA and worldwide, mostly for multinational and Fortune 500 corporations. He is a co-author of multiple publications and speaks at conferences about Internet, service-oriented architecture (SOA), and security solutions on System z. Foulques is an IBM and The Open Group Master Certified IT Specialist and a Certified Information Systems Security Professional (CISSP). He received a Master's degree in Computer Science and Engineering from Ensimag in France. He furthered his education at the State University of New York in Buffalo and at Stanford University in California, USA.

Christopher Dziekan is a seasoned Executive with 22 years of experience in business intelligence and analytics, having worked for enterprise BI companies including Crystal, Hyperion, Business Objects, Cognos®, and is now both the worldwide Business Analytics Cloud Portfolio Executive for IBM and the overall Executive responsible for the Office of Strategy for Business Analytics.

Michael Dziekan is a Senior Strategy Advisor for Information Management Strategy & Planning. Michael has been a contributing thought leader and pioneer of innovative world wide enterprise programs at IBM for the past five years, drawing on over 20 years of industry experience. Michael developed the IBM Information Management internal framework (NIMA). He also played a pivotal role in helping Cognos become a leader in Performance Management. Responsible for developing the Cognos global Standardization & Business Intelligence Competency Center (BICC) vision, Michael is recognized by the market, analysts, and press as a foremost visionary and subject matter expert

(SME) on the topic of operational business models. His published BICC research and studies with over 400 customers world wide across all geographies and industry markets enabled customers with winning conditions for accelerated and sustained success of their technology investments.

As spokesperson and evangelist to the broad dynamics of Information Management, his work with senior management and executives through world wide research ensures that our global customer market is well understood and aware of best practices and enabled to realize dramatic benefits from our technologies and solutions.

Andrea Greggo is a Cloud Computing Lead working in Market Strategy for the System z brand for the past two years. She has worked for IBM for 11 years in a variety of technical and business positions. She has an MBA and an MS in IT from Rensselaer Polytechnic Institute and a BS in Computer Science from Mount St Mary College.

Christian Hagen is a Senior IT Architect at IBM Global Business Services® in Germany. He has 18 years of experience in the IT industry. He has a diploma in physics from Christian Albrechts Universität Kiel. His areas of expertise are application development, data warehousing, data migration and architecture methodology.

Douglas Lin is a Senior Technical Sales Specialist at IBM Sales and Distribution in the United States. He has over 10 years of experience in the IT industry and has worked with customers in the financial markets, media, public, and computer services sectors. His areas of expertise are consultive sales, solution design, and product development. He has both a Bachelor and Master of Science degrees in Electrical Engineering from Cornell University.

James Machung is a Team Leader in IBM Global Technology Services in the United States. He has 15 years of experience in the IT industry, including 10 years with IBM. He is an IBM Certified DB2® Database Administrator whose areas expertise are database management, CRM, and business intelligence applications. He has additional expertise in data analysis and business analytics programming in the public health and life sciences sectors. He has a Master's degree from Pennsylvania State University.

Nicole Roik is a Senior IT Architect and works as an Industry Architect for IBM Sales and Distribution Insurance in Germany. She has experience in domestic and international service projects, which includes application development, IT architecture, and consulting. Her areas of expertise are Enterprise Architecture and Information Management. Nicole has a diploma in Economics and has worked for IBM for 12 years.

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author - all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

► Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

► Send your comments in an e-mail to:

  redbooks@us.ibm.com

► Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- ► Find us on Facebook:

  http://www.facebook.com/pages/IBM-Redbooks/178023492563?ref=ts

- ► Follow us on twitter:

  http://twitter.com/ibmredbooks

- ► Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- ► Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- ► Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

Part 1

Introduction to cloud computing

The next evolutionary stage of information technology (IT) manifests itself as cloud computing. Cloud computing represents a shift towards a more user-centric computing model where a layer of services sitting on top of an infrastructure decouples core business functions from the delivery of the underlying technology. IBM recognizes that there are multiple ways to deliver IT capabilities, which includes traditional software, hardware, and networking approaches, pre-integrated systems and appliances, and new breakthroughs that are provided as a service. In addition to the hardware and software elements, companies must also consider the cultural, funding, and business process changes that are involved in cloud computing. Typically a BI Competency Center (BICC) is a core ingredient to success in standardizing and driving an information-led business optimization transformation. To learn more about sustained success through operational efficiency, see Appendix B, "Competency centers: Sustained success through operational efficiency" on page 299.

There is a greater need for IT to help address business challenges. IT is expected to do more with less (reduce capital expenditures and operational expenses), help organizations reduce risk (ensure the right levels of security and

1

resiliency across all business data and processes), improve quality of services and deliver new services that help the business grow, and increase the ability to quickly deliver new services to capitalize on opportunities while containing costs and managing risk.

By itself, the new paradigm presents many opportunities, including improved cost efficiencies and rapid deployment of computing resources. Application of cloud computing to specific services, however, can open the door to even more possibilities. The IBM Smart Analytics Cloud harnesses cloud computing to business intelligence and analytics, aspiring to revolutionize information processing and decision making.

Part one of this book provides a framework for understanding the Smart Analytics Cloud and discusses the value of applying cloud computing towards business analytics. We dive deeper into the details of an IBM service offering later in this book. Key concepts and definitions are introduced to help you evaluate the benefits of deploying cloud-based analytics in your business.

1

IBM Smart Analytics Cloud

Our modern information environment is more complex than ever. Data volumes are burgeoning and coming in from all mediums, including blogs, podcasts, Wikis and tweets. The pace of everything is accelerated. Information must be analyzed, contextualized, and shaped for decision-making and right-timed action. At the same time, globalization demands better sharing of information not only with colleagues down the hall but also with those around the world. The IBM Smart Analytics Cloud responds to the challenges that are posed by the information explosion and flattening world, helping businesses seize the opportunity to gain a competitive advantage through business intelligence and analytics.

The IBM Smart Analytics Cloud is a service offering that enables the delivery of business intelligence and analytics at your location in a private cloud deployment. Its objective is to make businesses smarter, empowering organizations and enabling all employees, especially those closest to clients and suppliers, to make better decisions.

Transformation to cloud computing changes the economics of business intelligence and analytics. Rapid service provisioning times enables a variety of new analytic data management projects and business possibilities. Innovations and new technologies can be introduced in less time. Cloud computing fundamentally presents a more efficient and cost effective deployment model.

Change, however, is not without its challenges. Leaders must understand why and what change is necessary to best reshape their businesses to compete. IBM recognizes the potential obstacles to moving to a cloud computing model. In

3

many enterprises, lines of business often equate information with control and resist moving data into a more shared deployment model. To help businesses realize the benefits of cloud analytics, IBM focuses the Smart Analytics Cloud on business intelligence and analytics, allowing lines of business to control and manage data. Figure 1.1 shows how the Smart Analytics Cloud fits into a business intelligence and data warehousing reference architecture, which is documented in section 5.1, "Data warehouse environment" on page 48.
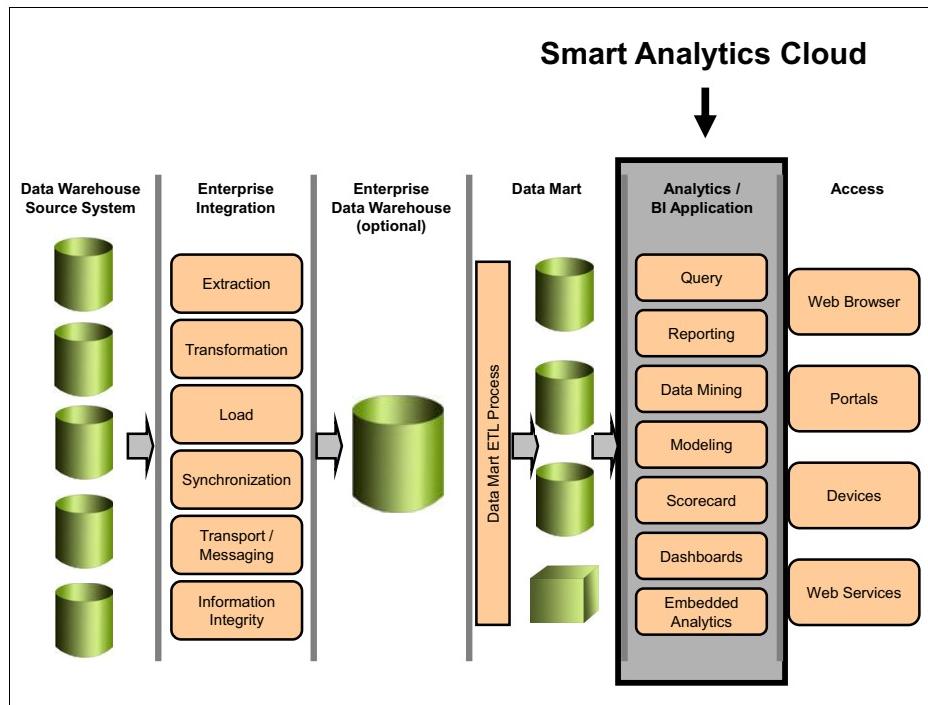
Figure 1-1   Where the Smart Analytics Cloud fits in

The Smart Analytics Cloud service offering provides a complete solution that enables a business to create a core approach to delivering business intelligence and analytics across an enterprise. IBM itself implemented a Smart Analytics Cloud that is known internally as *Blue Insight*. The internal private cloud supports 1 petabyte of data and helps more than 200,000 global employees make better decisions by providing them with real-time information about customers and suppliers, whether they are in offices or in the field. It serves as a template for the IBM customer-oriented service offering.

1.1 Evolution to cloud computing

While the economic downturn that began in 2007 has constrained budgets, it also stimulated transformation to cloud computing and quickened the adaptation of its key enabling technologies, such as virtualization. For clients looking to lower costs, perhaps by deferring capital expenditures or off loading non-core IT processes, cloud computing presents a way to do so while still providing services and deploying them quickly. Cloud computing will become even more widely used as technologies, such as virtualization, automation, and provisioning, mature.

Cloud computing is both a business delivery model and an infrastructure management methodology. The business delivery model provides you with a standard offering of services, such as business analytics that are easily accessed and rapidly provisioned. Steps, such as producing hardware, installing middleware and software, and provisioning networks, are dramatically simplified. The infrastructure management methodology is built on virtualized resources and provides better economics and increased ability to scale. It makes high volume, low cost analytics possible.

But not all clouds are created equal. Important attributes, such as location, ownership, access, targeted users and workload (application types), varies across an array of clouds, as shown in Figure 1-2.
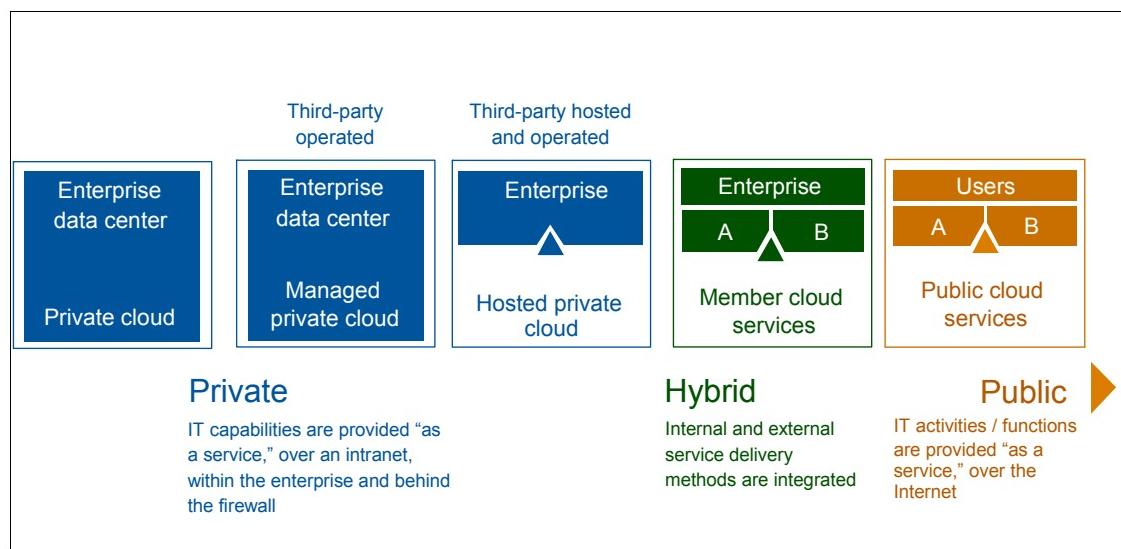
Figure 1-2   Attributes vary across an array of clouds

Private Clouds are dedicated to a company and can be *on-premise* or hosted by a trusted third party. Capabilities are owned by IT and access is through an internal network. Capabilities (hardware and software) are then dedicated to the individual company, leveraged across the multiple departments and lines of business. Many companies that are interested in cloud computing are ready for *Private Clouds*. However Public and *Community Clouds* are expected to grow significantly over the next three years.

For you to get the full economic benefits of cloud computing, you must be willing to virtualize your environment, standardize and automate your hardware and software, and share or pool infrastructure.

Interim steps can be taken to lower the cost of computing, but each step requires that you are willing to trade control of the infrastructure for cost savings. Creating a roadmap for cloud computing must be part of an IT optimization strategy:

1. The journey begins with *consolidation* to reduce infrastructure complexity, reduce staffing requirements, manage fewer things better, and lower operational costs.

2. A *virtualization strategy* is adopted to remove physical resource boundaries, increase hardware utilization and costs, and simplify deployments.

3. S*tandardization and automation* occur to unify your organization on a set of standard services that reduce deployment cycles, enable scaleability and flexible delivery, and increase the ability to reach a closer version of truth across the disparate data and disparate Business Intelligence (BI) silos.

IT Professionals must balance reward, risk, and control as they consider what type of cloud to deploy and what workload to place upon it. In today's data centers, IT has control (where the center is located, administrators have direct physical access, the audit process is clear and understood, and the internal security team is involved). In tomorrows Public Cloud, questions arise, including where the data is located, who has control and ownership, who backs it up, how resilient is the system, how do our auditors observe it, and how does our security team engage. However, the rewards increase with more reuse and economy with scale sharing of resources.

It is important to understand the workload that is being imposed on the cloud infrastructure and make smart decisions about which workloads are initially ideal for Private Clouds (database and application-oriented workloads emerge as most appropriate) because of the nature of the business process, data, and security requirements, and recognize that other workloads are fit for Public or Consortium-Style Clouds (Infrastructure and Collaborative workloads emerge as most appropriate)[1].

---

[1] Source: IBM Market Insights, Cloud Computing Research, July 2009.

The top private workloads, where database and application-oriented workloads emerge as most appropriate, are:

► Data mining, text mining, or other analytics
► Security
► Data warehouses or data marts
► Business continuity and disaster recovery
► Test environment infrastructure
► Long-term data archiving and preservation
► Traditional databases

The top public workloads, where infrastructure and collaborative workloads emerge as most appropriate, are:

► Audio/video/Internet conferencing
► Service help desk
► Infrastructure for training and demonstration
► WAN capacity and VOIP Infrastructure
► Desktop
► Test environment infrastructure

Is cloud computing just another buzz word for an existing computing model? Unlike earlier paradigms, such as grid computing that move workload to computer resources, the cloud model moves computer resources to the workload. Perhaps cloud computing is better explained by an analogy. Consider the transformation of the semiconductor industry. At one point, chip vendors all had fabrication plants. Today, however, there exists many fab-less companies that succeed by focusing on innovative chip design without the capital, operational expenses, and risks that are associated with owning a state-of-the-art fabrication plant. At the same time, the companies with fabrication lines lowered cost and risk by sharing their production resources among multiple customers. Cloud computing, similar to the fab analogy, separates the end user from the infrastructure. Just as chip designers can now specialize on innovative design, business analysts can specialize in analytics while leveraging underlying technologies that a cloud deployment provides.

Cloud computing, however, likely will not play out in exactly the same manner as in the semiconductor example because it spans a much larger scope than chip fabrication. While one might jump to the conclusion that there will eventually be a few large public clouds, what is more likely is a hybrid model of both public and private clouds.

Private clouds, sometimes referred to as *internal and secure clouds*, are client dedicated and have access and security defined by a client. Access is limited to client and partner networks, allowing for more control over service quality, privacy, and security. In general, private clouds are also restricted for use behind a company firewall and therefore have fewer security exposures.

Security is one of main reasons for selecting a private cloud model instead of a public cloud model. A business that has customer or sensitive data is concerned about having data in clouds. For a financial markets firm, a breach can result in significant costs and damages. Other reasons include availability, auditability, and guaranteeing service levels. For many enterprises, public clouds are not deemed to be reliable enough yet for specific workloads that are related to sensitive data.

A private cloud can offer a variety of services to multiple organizations. Figure 1-3 shows an example of an enterprise private cloud and several sample cloud-based service offerings, one of which is business intelligence and analytics. Figure 1-3 also shows how multiple organizations, such as human resources (HR), sales, and marketing, can all use the same set of services.
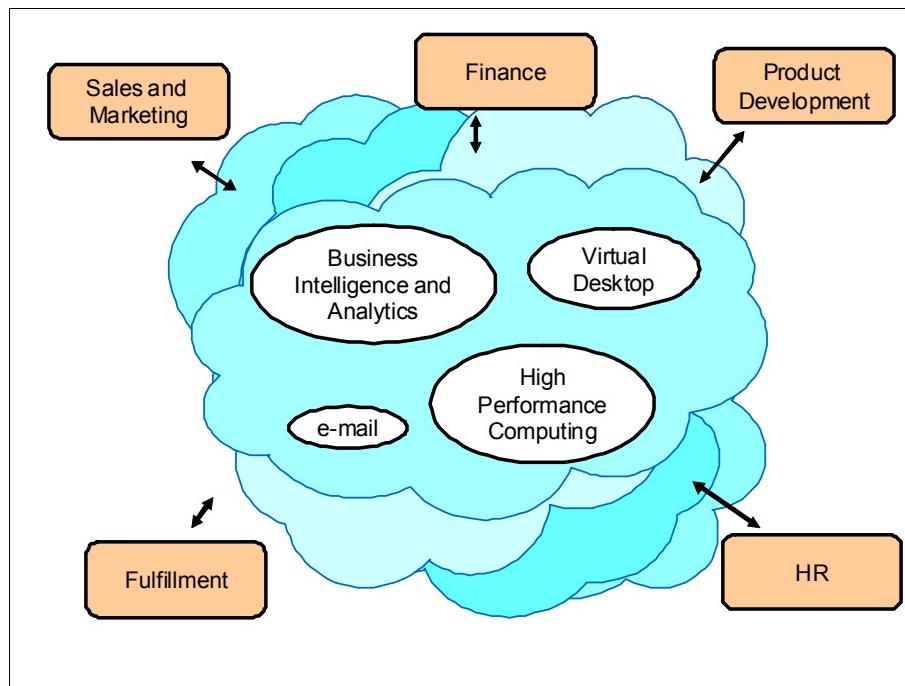
Figure 1-3   An Enterprise Private Cloud

1.2 Clouds and business analytics

In this section, we discuss optimizing business through the use of business analytics. We also discuss the IBM private cloud solution, Blue Insight, and how it

democratized information, providing access to a variety of client and market data regardless of where an employee sits in the company.

1.2.1 Business analytics and optimization

Organizations are shifting investments to leveraging information for Smarter Business Outcomes. Business leaders are telling us that to meet their goals for profitability, revenue, cost reduction, and risk management—especially in the current economy—they know they cannot continue to operate the way they have in the past. Simple business automation initiatives have only taken them so far, and they are realizing that through the better management and use of information—information that might already be at their disposal or easily gathered—they cannot only remove the blind spots that are keeping them from making informed decisions, but they can also achieve the next generation of efficiencies by providing precise, contextual analytics and insight at the point of impact.

Primary IT investments over the past two decades focused on automating business processes with the objectives of driving faster processing and reduced costs. This focus was driven by an *application agenda* to implement ERP and financial applications, supply chain management solutions, and call center and Customer relationship management (CRM) applications. However, these types of investments are no longer creating a sustainable competitive advantage for organizations.

As a result, over the past few years, we saw new initiatives increasingly focused on optimizing their business to drive a more sustainable competitive advantage in the marketplace while reducing costs. This focus means moving from just leveraging ERP and financial applications to providing increased financial risk insight for better business decision and moving from just managing your supply chain to enabling more dynamic demand planning and moving from just managing your call center and customer relationships to providing increased insight to improve customer service to drive greater profitability from your customers.

These new initiatives are all dependent on information and having an information agenda in place. To do this, a company must have:

► A strategy: Establish an information driven strategy and objectives to enable business priorities

► A roadmap: Accelerate information intensive projects that are aligned with the strategy to speed both short and long-term returns on investment

► An information infrastructure: Deploy open and agile technology (including cloud computing) and use existing information assets for speed and flexibility

► A defined governance plan: Discover and design trusted information with unified tools and expertise to sustain competitive advantage over time.

Just as not all clouds are created equal, not all business intelligence offerings in the industry are created equal. There are at least five key differentiators for IBM business analytics:

► Succeed Today

A complete Performance Management system versus a suite of un-integrated products that are built on various technologies

► Protect Tomorrow

A Performance Management System that is open, modular, and extensible versus a history of unfulfilled promises and detours

► Innovation without Risk

Innovations that are progressive, pragmatic, and purposeful versus un-integrated product introductions that disrupt the clients' business

► Analytics for Everyone, Everywhere

Access to powerful analytics are needed for all business users across all data sources, platforms, and the full spectrum of analysis

► Leadership and Expertise

The industry's only practice dedicated to Business Analytics and Optimization and the delivery of customer and industry best practices through blueprints and applications

IBM delivers the full range of integrated capabilities that address the critical questions that decision-makers must answer. What really sets IBM apart is not just this full range of capabilities, but does it do so in a way that decision makers see a complete, consistent, and trusted view of information. One such offering that provides a high level of expertise is the IBM Smart Analytics Cloud service, which you can read more about in section 3.3, "The IBM Smart Analytics Cloud" on page 24.

An organization consists of people with various skills and roles all trying to pull in the same direction with the goal of optimizing business performance. Each of these people require multiple levels of information and detail to make decisions that impact performance. IBM offers the complete range of integrated Business Analytics capabilities to address this full range of user needs.

Using highly visual scorecards, dashboards, reports, and real-time activity monitoring, decision makers gain immediate insights regarding the health of the business and can understand what is happening in their area of the business.

Analyzing trends, statistics, correlation, and context, decision makers can understand what leads to the best outcomes and discover why things are on or off track. Knowing what is likely to happen equips decision-makers with the foresight that they need to intervene. Simulation through predictive modeling and what-if analysis enables decision makers to predict and act and change the course to improve the outcomes. Financial and operational planning and budgeting and forecasting puts resources in the right place and sets targets for those allocations.

Everyone in the organization can be confidant in a common, consistent, and trusted data. IBM allows you to pull data from a range of systems and makes it easier to turn this data into information. The knowledge level does not matter, everyone can consume the information in a manner that is relevant to them.

The right information, in the right way, to the right people, at the right time, leads to optimized decision making.

1.2.2 The IBM internal business analytics cloud

The IBM internal Blue Insight service is a private cloud that enables IBM to standardize on a single BI Solution (IBM Cognos 8) across the enterprise. Coupling the private cloud model with the System z platform ensures top-notch security and availability for the IBM business intelligence and analytics service. The in-house cloud supports over 200,000 knowledge workers globally who require access to business intelligence and analytics to do their job. It was built to address a key enterprise problem, which is how not only to collect data but how to make it widely available for use. The objective follows a trend in business intelligence and analytics that is to go beyond just pulling historical data. Part of the challenge and the potential opportunity for exploiting analytics is to push the right data, at the right time, to end users.

As an example, consider a scenario where a sales manager is reviewing sales reports. The reports show that sales of items on promotion increased 10%. At first glance, one might approve funding for a similar but larger promotion. More information might be needed though. Using innovative business intelligence and analytics techniques, the sales manager might be informed that no items other than the discounted items were sold. In that case, the sales manager does not approve a larger promotion because it did not drive purchases of related full-priced items and actually reduced margins.

Consider also a credit card company. Such a company can better match product offerings with customers based on credit risk, usage, and other characteristics. Today, there is also more of a focus not on just reports of what already happened but how it happened and why. So not only can a company look at customer credit

history but it can also rank individuals by their likelihood of making future payments.

Significant value is achieved by applying business intelligence and analytics, which can help executives make more informative decisions by providing fact-based answers to fundamental business questions, such as:

► Who will be our most profitable customers?

► What will be the impact on profits when introducing a new product line?

► How would a price change influence the behavior of various customer segments?

► Do recent purchasing patterns represent the start of a long-term trend, cyclical behavior, or just a short-term aberration?

Implementing a business analytics solution like the IBM Smart Analytics Cloud can improve business' ability to answer these and other questions.

2

Building a Smart Analytics Cloud

Underlying technology is essentially transparent to the end user in the cloud computing model. The selection of components is left to the cloud service provider and largely depends on functionality, service levels, and costs. For an enterprise service, it is also imperative to address the key concerns of security and availability.

To best meet those requirements, the IBM Smart Analytics Cloud is built upon two key building blocks:

► IBM Cognos 8 Business Intelligence software (IBM Cognos 8 BI)
► IBM System z platform

IBM Cognos 8 BI is a proven and powerful product that provides a complete range of business intelligence and analytics capabilities, including reporting, analysis, scorecards, and dashboards. As shown in Figure 2-1 on page 14, IBM Cognos 8 BI services can be accessed through various ways, including Web 2.0 interfaces, a desktop office product, and smart mobile devices. System z provides industry-leading virtualization, disaster recovery, security, resiliency, and scalability. These are the same building blocks that are used in the IBM internal Blue Insight private cloud. Blue Insight runs on a System z and uses cryptographic hardware accelerators to handle up to 10,000 secure transactions per second.
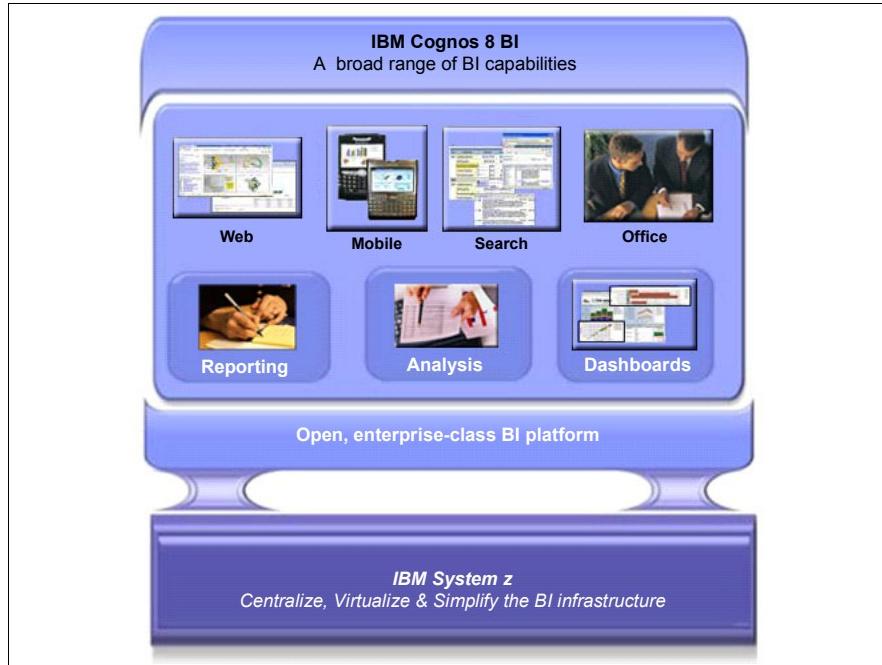
Figure 2-1   Core components of the Smart Analytics Cloud

2.1 Why System z

Cloud computing might be a relatively recent term, but key elements, such as the concept of virtualization and timesharing, have been around for decades. Developed and enhanced over many years, System z showcases the industry's most robust and mature virtualization environment. In addition, many of System z's core strengths are now finding direct application to a private cloud environment. Availability, scalability, and security are key technologies in a private cloud and are also strengths of the System z platform.

In addition to increased capabilities, another key reason for selecting System z is decreased cost. If it seems paradoxical to associate System z with lower cost, take a step back and consider the larger total cost of ownership perspective. System z environments have the potential to consume less energy than distributed environments, decrease software license costs, lower network equipment costs, reduce real estate requirements, and require fewer administrators. The economic downturn fueled a wave of consolidation and, of multiple consolidation alternatives, the combination of Linux and System z is one of the most compelling. The primary reason System z is a strong consolidation

platform is that its processor is architected to run effectively at near 100% utilization. Distributed architectures operate more effectively at lower utilization rates. Therefore, System z processors can fundamentally handle more workload per core, even at equal or lower clock frequencies than other processors. Platform selection is important. The Smart Analytics Cloud leverages System z as an underlying hardware technology to provide a private cloud service that is resilient, scalable, and secure.

See sections 5.3.2, "Description of the building blocks" on page 56 and 5.4, "Operational overview" on page 59, for an in depth discussion about the System z platform.

2.2 Cloud management

While the Smart Analytics Cloud is built using two core components, IBM Cognos 8 BI and System z, these pieces must be tied together and presented as a fluid and responsive computing service. Management functionality, which we refer to as cloud management in this book, is the glue that binds the cloud components together. Cloud management encompasses a set of tools, processes, and capabilities that provides services, for example, business intelligence and analytics services, to an end user.

The goal of cloud management is to reduce complexity through automation, business workflow, and resource abstraction, for example, a user wants an analytics environment to test a marketing model. The user, also known as the service requester, browses through an IT service catalog and submits a request for a test environment. A service manager approves and the cloud administrator sets up the remaining tasks. The steps can be completed in minutes instead of months and, just as important, are transparent to the end user. Cloud management streamlines processes and can save weeks or months of time that it often takes to procure and configure hardware, operating systems, applications, networks, and storage.

Multiple components and challenges must be considered when deciding to proceed with a Smart Analytics or other cloud solution. The key aspects include automation, provisioning, monitoring, security, capacity planning, and onboarding. Onboarding is the process of installing and configuring the operating system and additional software on servers to meet end user requirements. Manual onboarding is a time and labor consuming and error-prone process. Automation of such processes presents an opportunity to improve efficiency, reduce errors, and get more value from a private cloud deployment.

Onboarding: Onboarding is the process of installing and configuring the operating system and additional software on servers to meet end user requirements.

Management tools and processes are vital to operating more complex virtualized cloud environments. Virtual resources must be managed so that virtual sprawl does not occur. You do not want to invest in virtualization only to learn that the management overhead offsets the savings. Proper architecture and implementation of cloud management is critical for the success of a private cloud.

Actual architecture and implementation of cloud management is flexible and can be built to fit specific needs. Cloud management typically begins with centralization and standardization tools and processes and progresses from there. The objective is to get people the information that they need to learn, react, and make better decisions. The evolution to cloud computing is just a start. The build out of services, such as Smart Analytics, better equips a business and creates more opportunities for revolutionary innovation.

Part 2

Business

In this part, we introduce the business objectives for analytics and business intelligence. We analyze what aspects of analytics and business intelligence can be provided efficiently using a cloud compared to the approaches that are used today without a cloud. Putting analytics and business intelligence together, we show the specifics of a Smart Analytics Cloud, including the considerations that we made for providing the service that is included in the cloud. Additionally we give an overview of the service offering that IBM has available to support our clients to implement a cloud in their enterprise.

3

Business objectives

As in the past, adapting to change is critical to building and sustaining a competitive advantage. Change, while it is reflected by disruptions in the marketplace, also leads to new ways to do business. Instead of being slowed down by the influx of information, organizations can use new approaches to take advantage of it. By using innovations in technology, our ability to analyze information takes a leap forward. Where organizations before relied only on intuition, they can now use business intelligence and analytics for fact-based decision making and answering such questions as how to design price offerings, what markets to target for new services, and how to reduce risk exposure.

In this part of the book, we discuss the value of applying cloud computing to business intelligence and analytics. In the first section, we present the reasons for business intelligence and analytics. In the second section, we discuss what cloud computing is and why it is applicable, and in the third section, we discuss the specific values of the Smart Analytics Cloud.

3.1 A case for Smart Analytics

Information and the technology that drives it continues to evolve and, as it does, change the way we live and work. Business intelligence and analytics aim to improve our decision making by translating volumes of data into valuable insight. With the capability to better analyze and comprehend data, employees from C-level executives to first-line staff are empowered to present fact-based analysis and influence potentially game changing decisions. Organizations can challenge the status quo and take bold steps to improve performance. The goal, ultimately, is to get the right information to the right people at the right time. The challenge is to transform an organization's capabilities to effectively provide business intelligence and analytics services.

In addition to processing data at increasing rates, another challenge is turning data into relevant information. Information flows much faster today. The capability to access relevant data and better anticipate future outcomes has significant value. Such forward looking abilities can help an organization weigh trade-offs and make better decisions about future pursuits. Business intelligence and analytics technology can help to turn growing amounts of information into insight.

The scope of computer usage has expanded. When information technology was introduced, businesses improved by simply using computers to automate repeatable tasks, such as forms processing. Today's innovations can sort through vast amounts of data and transform that information into intuitive reports and scorecards. Think about the potential impacts. Among many other opportunities, you can exploit innovative technology to make better decisions about where to market new productions and services or to use medical information to provide better medication or to improve the traffic situation in cities.

The fundamental concepts behind business intelligence and analytics are not new. Quantitative and analytics methods were used in businesses, such as financial markets trading, for some time. What changed today is the maturity of the technology and tools that provide the business intelligence and analytics services. Current tools, such as IBM Cognos 8 BI, have intuitive user interfaces and, coupled with a cloud computing deployment strategy, can make business intelligence and analytics tools easily accessible across the entire organization and not just dedicated user groups.

Picture an organization where pertinent information, perhaps customer or inventory data, is shared effectively and in a timely fashion. Individual employees, instead of each storing similar information in their own spreadsheets, can generate customized reports using a common service. Through that and similar scenarios, businesses can take advantage of business intelligence and analytics

methodologies to improve information sharing and bring fact based analytics to mainstream business.

Technology, however, is only the beginning. How one uses the technology provides the even greater value add. Let us say that a business implemented an analytics software and hardware solution. In a typical model, a business analyst submits a request to an IT service group to analyze and provide a report. The IT staff, while being highly skilled in technology, might have knowledge of the businesses that they serve but are not the experts. Efficiencies are gained and better insights gleaned if the business analyst can analyze data without needing deep technical knowledge. Decision-making speed improves. Business analysts have direct access to analysis and information and do not have to wait for IT resources to become available. An unneeded level of communication is removed. By moving tasks, such as creating reports from IT service groups, those resources are freed up to develop new innovations.

3.2 Cloud computing

Before transforming your IT capabilities to effectively provide business intelligence and analytics services, it is important to think about what it might look like. Instead of reinventing the wheel, you might model an IT service based on successful existing service models, such as water or electricity utilities. Public utilities can service many consumers while standardizing and centralizing delivery. They use economies of scale to provide competitive pricing and additional value. After some further consideration, you might consider these operational requirements: scalable, resilient, elastic, automated, and standardized. You can also envision an environment where the end users have a simple method to request services and that there are easy processes for adding, maintaining, and sun-setting services.

With so many different aspects, where do we start when developing a business intelligence and analytics service? First, let us consider the operational parts of a service and determine if they, like a public utility, are good candidates for centralization.

The four key operational parts of a service are:
► Hardware
► Software
► Data
► Business applications

Hardware infrastructure can be managed centrally by an organization or locally by lines of business. The local management approach resulted in a proliferation

of server farms, which has reversed because organizations are leveraging virtualization and want to benefit from economies of scale.

Business Intelligence software and middleware can either be purchased on a department-based level or on an enterprise level. Over time, departments might develop different preferences and skill sets, but such preferences are often outweighed by savings that can be achieved when negotiating for either a larger number of licenses for one product or a smaller set of enterprise-wide software. Further costs reductions are seen in administration and maintenance costs too.

Centralization of data, however, receives strong objections by departments. Responsibility for data is a sensitive topic. Lines of business must be in control of the information that is important for the business that they are responsible for and therefore want to manage data themselves, which leaves them enough flexibility to react to changes in the market.

Business applications provide strategic value to each line of business. While the underlying software and middleware can be centralized, it does not make sense to use the same approach for the value differentiating end-user application.

Of the operational components, the hardware and software infrastructure components lend themselves to a centralized approach. Centralization and standardization of that infrastructure is known as cloud computing. While cloud computing signals a shift from a distributed to a centralized mindset, there is real value in such a change. Lines of business can focus on what they feel is more important, their data and business applications, and can obtain the reliable infrastructure from the specialized cloud provider.

The term cloud computing is used in different ways. Its usage, however, does have common themes. On one hand, cloud computing is an infrastructure and services methodology. On the other hand, it is also a user experience and business model.

Cloud Computing is:

▶ An infrastructure management and services delivery methodology: Cloud computing is a way to manage large numbers of highly virtualized resources such that, from a management perspective, they resemble a single large resource, which can then be used to deliver services with elastic scaling.

▶ A user experience and a business model: Cloud computing is an emerging style of IT delivery in which applications, data, and IT resources are rapidly provisioned and provided as standardized offerings to users over the Internet in a flexible pricing model.

The cloud meets key requirements, which are:

► Scalability: Can increase capacity without impacting functionality.

► Resiliency: Allows applications to continue functioning even when underlying components fail.

► Elasticity: Can add or change functionality without changing or disturbing existing functionality.

► Automation, standardization: Adding resources in a standardized way and, wherever possible, in an automated way.

► Service life cycle support: Setting up new infrastructure and software, maintaining it, and sunsetting it.

► Self Service: Provides an easy-to-use interface that allows end users, who might not have deep technical skills, to request new resources.

Cloud computing is not just an improvement in data center infrastructure, but it is also a user experience and business model. In a cloud deployment, the end user sees standard offerings of services that are easily accessed and rapidly provisioned. Figure 3-1 depicts a cloud and its basic components.
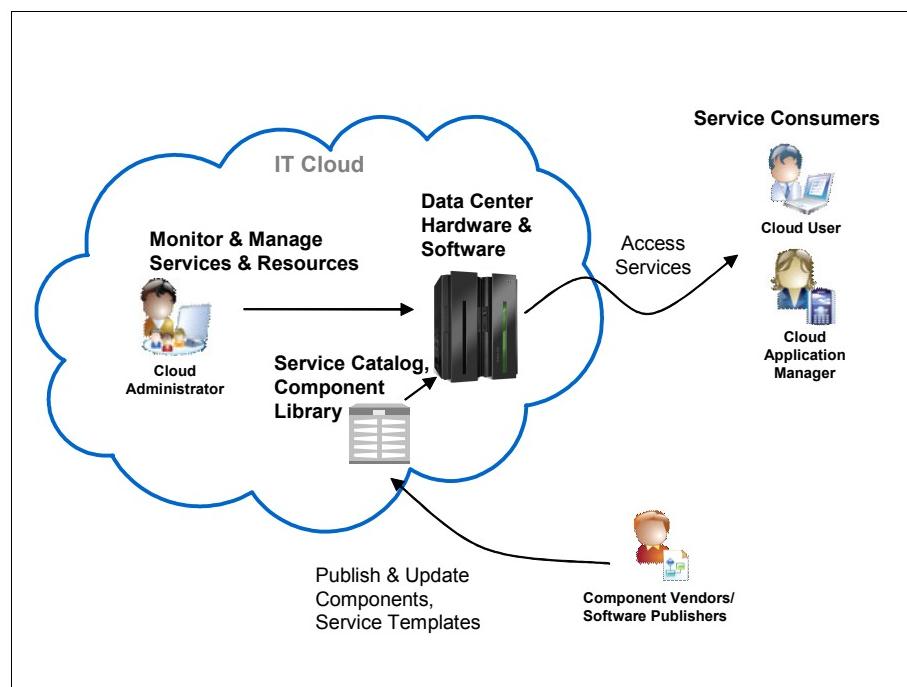
Figure 3-1   Components of a cloud

The basic internal components of a cloud are data center infrastructure, a service catalog, and a component library. Data center infrastructure includes hardware, such as a System z, software such as IBM Cognos 8 BI, and middleware like DB2. The component library encompasses the hardware, software, and service components that are required to deliver services. The software catalog lists the services that are provided to the customers, for example, the service catalog can include installation of a Linux guest, licenses for Cognos, or even a complete Smart Analytics Cloud service.

The key roles in the cloud are service consumers, administrators, software publishers, and component vendors. Service consumers make requests through access services that have a standard user interface. Cloud administrators monitor and manage the services and resources that are delivered. Software publishers might be internal departments that develop customized services. Component vendors, such as IBM, can also offer services, such as the Smart Analytics Cloud.

In discussing cloud computing, it is also important to distinguish between types of clouds. While there is variation in the naming, there exist predominantly to classifications, public and private[1].

► A *public cloud* is owned and managed by a service provider and access is through subscription. It offers a set of standardized business processes, application, and infrastructure services on a price-per-use basis. Advantages of a public cloud include standardization, capital preservation, flexibility, and a shorter time to deploy applications.

► A *private cloud* is accessible only through your company and your partner network. It provides more ability to customize, drives efficiency, and retains the ability to standardize and implement best practices. Other advantages are that the levels of availability, resiliency, security, and privacy are determined on an enterprise level independently from an external provider.

For many organizations, public clouds are not secure or reliable enough. Private clouds provide increased flexibility and are used for enterprise class solutions.

3.3 The IBM Smart Analytics Cloud

Transformation of your business intelligence and analytics capability can lead to game changing decisions, help anticipate future outcomes, and empower employees. The IBM Smart Analytics Cloud service offering is designed to overcome challenges to change and enable customers to successfully ramp up their business intelligence and analytics capabilities.

---

[1] See the IBM Redpaper™ publication, REDP-4553-00 p. 2

> **The IBM Smart Analytics Cloud is:**
>
> A service offering that enables the delivery of business intelligence and analytics at a customer location in a private cloud deployment. Its objective is to make businesses smarter, empowering organizations and enabling all employees, especially those closest to customers and suppliers, to make better decisions.

To both benefit from and prove the value of its service offering, IBM executed the Smart Analytics Cloud internally. The in-house solution is referred to as Blue Insight and its objective has been to transform IBM by developing an enterprise-wide business intelligence and analytics strategy that leverages:

► Common services, infrastructure, knowledge, and processes for the Analytics and Business Intelligence domain.

► Centralized infrastructure that empowers organizations to utilize their domain expertise for transformation initiatives.

Results from executing Blue Insight resulted in many benefits, including:

► Consolidated business intelligence and software product sets
► Expanded amounts of data sources that a single end user can tap into
► Ability to serve over 200,000 users
► Cost savings from hardware, software, and operational efficiencies
► Increased elasticity

Change at IBM did not come with out its challenges. Different lines of business had their own business intelligence methodologies with different software and hardware products. There was resistance to centralization, not only of data, but to other aspects of the service. At the end though, IBM realized significant value. Going forward, IBM is ready to help its customers learn from its own experience and enable them to reap the benefits of a cloud-based business intelligence and analytics.

The IBM Smart Analytics Cloud drastically reduces the number of departmental solutions to a single BI environment that can support vast numbers of users across the lines of business. Additionally, a Smart Analytics Cloud:

► Improves standardization by introducing single points of control for departmental business processes, corporate security, and compliance.

► More effectively uses skilled resources to support a common business intelligence and analytics delivery tools.

- Reduces the capital and operating expenses that are needed to support enterprise-wide services.
- Supports a self service approach to dispensing business intelligence and analytics services that reduces the time, resources, and costs for delivering services to new divisions, departments, and users.

With the Smart Analytics Cloud, IBM helps its clients to create a business intelligence and analytics service. Like IBM, you can expect to see positive results from a cloud computing, such as cost savings, the ability to support vast numbers of users, and simplification of the software product set. More efficient processes improve accessibility and enable service consumers from across an enterprise to tap into intelligence and analytics capabilities. Lines of business can use common analytics services to sort through information and make better fact-based decisions. Cloud based business intelligence and analytics offers significant potential. After it is enabled, such capabilities present opportunities that did not exist before.

The next chapter describes the service that is included in the Smart Analytics Cloud. The combination of the requirements given in this chapter and the next are the basis for the architecture described in Chapter 5, "Architecture overview" on page 47, Chapter 6, "Functional architecture" on page 63, and Chapter 7, "Operational architecture" on page 73.

Scope of the Smart Analytics Cloud

A Smart Analytics Cloud is not only the installation of hardware and software. There is more to a cloud than just the components. It is a service that is offered, including availability times, response times, operations, maintenance, and release and capacity management.

When making a cloud available, the provider must describe what the end-user (or service requester as this role is called in this publication) can expect when requesting this service.

This chapter shows the considerations that a cloud provider must take into account when delivering the service of a cloud. The specifics and qualities of the service, given as examples in this chapter, determine how the cloud is set up and which hardware, software, and service components are used.

The first section of this chapter describes the scope that is covered by the service of the Smart Analytics Cloud. The second section shows the scope that IBM offers to implement this service. The specifics that we use here are only an example and might be different for other organizations.

4.1 Scope of the Smart Analytics Cloud

The Smart Analytics Cloud provider must define what services, infrastructure, and software are to be included in the cloud. The decisions about which software and hardware to use depends, to some extent, on the service that is offered with the cloud. It depends on non-functional requirements, such as availability times, response time, or number of users. In this section, we discuss what the provider must consider and the decisions that must be made to define the service as granularly as possible. We also give examples about why we chose the specifics in our lab environment.

The service requester is responsible for the business intelligence application that is added to the cloud. The application consists of reports and, if applicable, data. The application has certain user groups that need to must be added also, such as Cloud Users and Cloud Power Users.

After the specifics are decided and documented they can serve as the basis of an agreement between the Smart Analytics Cloud service provider and the service requester. It can be included in a document of understanding between these two parties.

4.1.1 Offered scope

First the service provider must determine the scope to offer. Considerations include the offered components, the available environments, and whether support is included and at what levels. The service provider must define how to proceed when a various software levels are required by the requester. Example 4-1 shows considerations of scope.

Example 4-1   Scope of the Smart Analytics Cloud in our lab environment

```
Common infrastructure for development, test, and production for the
latest production worthy Cognos version (production worthy assessment
is done by the Smart Analytics Cloud provider)

The Smart Analytics Cloud provider will deliver an infrastructure and
provide operations support services for a shared Cognos BI production
environment. This will also include Cognos Product Support for Level 3
issues.

The hosting environment will be used to deliver report content only. It
is not meant to be used as a mechanism for data delivery. For example,
although it may be possible to deliver 100,000 rows of CSV data, this
will not be permitted.
```

The Smart Analytics Cloud provider will provide content and security administration as it pertains to deployment into the development, test, and production environment. This includes promoting content from development to test to production, creating new folders, applying user group security to folders, creating and applying user group security to data sources, connections and sign-ons.

In the event that an service requester "must" move to a higher level of SW than is currently available in production, the Smart Analytics Cloud provider can accommodate a new pilot and production environment. The requester will bear the cost of the installation and the new production environment until the general production environment is moved to that level at which time the requester will go back to a per named user charge.

4.1.2 Access to specialized functions

Having set the scope of the Smart Analytics Cloud in 4.1.1, "Offered scope" on page 28, the cloud provider must now define which user groups or Cognos roles will be allowed to use which Cognos components. In addition, they must determine the level at which the user groups for each service requester are managed. Example 4-2 shows sample boundaries that you might want to set.

Example 4-2   Sample boundaries statements

Service requesters will use a standard Cognos Connect interface

Service requesters will use the Cognos Viewer to process interactive reports and view report output versions in public folders.

The service requesters will own their user groups and will be responsible for performing group creation and administration. The service requesters will maintain the list of the user group members.

Scheduled reporting to Public Folders is limited to administrators or an approved/trained application administrator.
Method for requesting a report schedule would be via a defined functional user ID. Deletion/expiration of report output will be based on age and business justification. Report specifications will not be deleted unless aged over two years or no activity for six months

Query Studio and Report Studio access to Cloud Power Users will be allowed for a limited user population. Packages used for Query Studio

```
and Report Studio requires processing limits. Saves are to My Folders,
not Public Folders; Interactive report must process within 10 minutes.
Batch/Scheduled reports must process within 30 minutes
```

4.1.3 Governance

The service provider must define the governance procedures:

► How will a test environment be moved to production, and what tools will be used?

► Should it be possible that existing report structures are given to other departments that have the same requirements?

In our lab environment, we decided to use a tool that initiates changes and allows tracking, to a certain extent. The functional and implementation details of this tool are documented in 6.2, "Onboarding application" on page 66, and Chapter 12, "Onboarding application" on page 211. The process of onboarding and the involved roles is described in Chapter 16, "Onboarding" on page 253.

Example 4-3 shows a sample that you might want to perform using an onboarding application. In our sample lab environment, we allowed existing report structures to be used by other user groups; however, this is largely dependent on the type of user groups and to what extent reuse is put into practice.

Example 4-3   Considerations for governance implementation

```
Service requesters will initiate promotion to the next stage via the
onboarding process and the services provided in the onboarding
application.

Promotion turnaround time will be generally less than 48 hours,
assuming all standard documentation has been provided.

Promotion requests will be managed through the onboarding process and
the onboarding application.

The Smart Analytics Cloud provider will govern all reports running in
any development or test environment to preserve service to other
Service requesters.
```

4.1.4 Support

The service provider must decide about the type of support that each environment receives. Distinctions between the production, development, and test environments must be made concerning, for example, the response times during week days compared to the weekend. How will incidents and problems be reported? What severity levels are available, and are they handled differently?

Example 4-4 shows several of the decisions that you must make about how a cloud solution might be supported. These decisions also influence the choices that are made in the systems management architecture. When providing comprehensive support for the Smart Analytics Cloud, incident management, change management, and monitoring must be available. While incident and change management are not described in this IBM Redbooks publication, monitoring architecture and its components are documented in 8.2, "OSS: Monitoring and event management" on page 107 and Chapter 18, "Monitoring" on page 269.

Example 4-4   Decisions for providing application support

```
24x7 support for production severity 1 issues

24x5 support or development and test issues, with weekend call out
Any problems identified in the development, test, or production
environment will be reported via a defined process.

Premium support for Cognos, including Level 3 Support when deemed
necessary by the Smart Analytics Cloud provider.

The Smart Analytics Cloud provider will give operational support for
development, test support for all environments with access to Cognos
Connection, Report Studio, Analysis Studio, Query Studio 24/5; 24/5
support being from 9 PM Sunday to 5 PM Friday.

The Smart Analytics Cloud provider will provide operations support as
part of the standard service.

Development and test problems will be reported per the common
operations process.
The Smart Analytics Cloud provider will offer Level 1 and Level 2
support including opening problem management reports (PMRs) with Cognos
as is needed for initiation of Level 3 support.

On-call support available via common help desk for severity 1 issues.
```

```
Development, debugging, or coding of reports or metadata modeling is
not included as support.

Severity 1 issues must be submitted via the common processes.
Problem determination will be performed for severity 1 problems and the
disposition will be documented and shared with service requesters if
the incident was a Cognos product related issue.
```

4.1.5 Security

Because a wide range of user groups use the Cognos Services with sensitive
data, the cloud provider must make sure that appropriate security levels apply to
the Smart Analytics Cloud. Example 4-5 provides these security considerations.

Example 4-5   Security considerations

```
Reporting data sources are internal that at a minimum supports DB2
server encrypt authentication.

Network data sources will need to be moved to the according network
segment in order to be accessible.

Service requesters are solely responsible for the security access
management for the content they publish.
```

4.1.6 Folder handling and sizing limitations

When many user groups use the same environment, conventions are necessary
to define: where each user group stores their content without disturbing other
user groups, how long certain content is stored and backed up, what folders will
be used, and so on.

Example 4-6 gives a sample of conventions that can be made for folder locations,
available disk space, and usage restrictions.

Example 4-6   folder conventions

```
Service requesters using scheduled report views will output content to
My Folders. Each user has their own My Folders area for personal use.
No scheduled reports will publish to Public Folders unless approved by
the Smart Analytics Cloud provider.

Report output content in My Folders will be deleted after 13 weeks for
quarter over quarter reports.
```

My Folder content is limited to 50MB for each individual user, and
usage over 50MB will be investigated to determine if any action is
required to ensure environment stability and performance

My Folders is not to be used as an archive facility. If archiving is
required it should be provided by the data source.

Result set sizes limited to 5K rows such that processing times and
server/temp space usage are acceptable.

The Smart Analytics Cloud provides a common directory for placing image
content required for service requesters reports.

4.1.7 Connectivity and application integration

The service provider must consider the connectivity that is supplied to other
applications. These considerations include the responsibility for data quality,
sources and custom landing pages, products used for cube creation, and cube
maintenance. Example 4-7 shows considerations for connectivity and application
integration.

Example 4-7   Connectivity and integration of other applications

Development and maintanance of custom landing pages lie in the
responsibility of the service requester.

Preferred cube creation will use IBM InfoSphere Warehouse cubing.

If Cognos is used for cube creation the service requesters will be
responsible to build the cube. Notification of cube replacement will be
the responsibility of the service requesters via Event Studio. The
Smart Analytics Cloud provider will host the cube for production
reporting within the production environment. Cubes larger than 500 MB
will require an exception from the Smart Analytics Cloud provider.

Service requesters are responsible for availability of data sources and
the quality of the data content.

4.1.8 Response times

Because it is difficult to guarantee response times for applications from
end-to-end, the service provider must decide whether they want to define a

service level for this. In our lab environment, we did not commit to response times; therefore, the statements we make about this are shown in Example 4-8.

Example 4-8   Example response time commitment

```
There is no response time commitment, given that service requesters may
opt to create their own models and author their own reports. However,
operations will identify and engage the required personnel as is
necessary to rectify issues that are environmental in nature.

There is no performance guarantee with regard to response time or job
processing time. The Smart Analytics Cloud provider will not provide
load testing or performance testing.
```

4.1.9 Billing

Providing a cloud allows the cloud provider to offer a service that is based on consumption of resources; therefore, which services are included in the service must be defined, for example, what type of support, when billing starts, and in which cycles invoices are raised.

Example 4-9 provides sample decisions that you might want to make.

Example 4-9   Example billing documentation

```
Service requesters initially will pay for the number of development and
test users when they board the cloud: Service requesters will incur,
per named user, charges defined in their user groups when they board
into production, or 3 months after the date they boarded into the
development environment, whichever comes first.

Service requesters will be audited for usage and charged according to
the number of named users.

Service requesters may be audited monthly for named user detail to
confirm the number of end-users listed in the service requesters user
groups.

After hours support will be provided as bill-to-actual costs to the
service requesters requesting the support. After hours support is
defined as requiring weekend development or test environment support.
Billing will be a net add to the quarterly bill based on actual charges
for the event.
```

In the event a service requester on the Smart Analytics Cloud
production environment requests the provider operations support team to
be available on weekends or a holiday, the requester will be billed
actual cost for the support.

Call-out support is for server down or when Cognos Connection is down.
For development and test support, required weekend callout charges will
be billed based on the actual charges to the service requester.

Support required by a service requester outside of the standard Smart
Analytics Cloud support service hours (week-ends) will require at least
five working days notice so that special staffing arrangements can be
made. Off-hours support will be bill-to-actual. There may be uplift
charges for weekend or holiday coverage.

Billing will occur quarterly for infrastructure and operations based on
unique named users, Bills will be aggregated and sent to the service
requester.

Billing for development (modeling/reporting) will be based on a
document of understanding with the organization delivering the
resource(s).

4.1.10 Maintenance times, planned outages, and freeze periods

The architecture of the Smart Analytics Cloud, as defined in Part 3, "Architecture"
on page 45, heavily depends on the availability requirements. The service
provider must decide when or whether downtimes can be scheduled, and in what
time frame they must be announced up front, or whether the company's freeze
periods also apply to the Smart Analytics Cloud.

Example 4-10 shows several of the decisions you might make.

Example 4-10   Maintenance times and planned outages

Scheduled down time for development and test will be Saturday 9:00 PM
to Sunday, 9:00 AM.

Updates and System changes to the production server requiring changes
to the operating system platform supporting Cognos Version 8 will be
completed during the Saturday evening change management window, which
is weekly from Saturday, 09:00pm until Sunday, 9:00am. Occasionally,
this maintenance window may be extended due to Cognos system
maintenance.

The development and test system may be unavailable during the following
times: System changes affecting Linux operating system or System z
resources (memory or CPU) will be done during normal business hours
(Monday — Friday, 8:00am — 5pm). These outages will be scheduled five
business days in advance, when possible. Severity 1 issues will be
handled immediately.

The development and test system may be unavailable during the following
times: Cognos services are recycled weekly on Sunday evening. Service
is unavailable from 9pm until 10:30pm each Sunday.

The development and test environments will be available during the
quarter end freeze period.

Deployments to production environments will be subject to the corporate
change freezes.

4.1.11 Communication

The purpose of a communication plan is to facilitate communication and mitigate
communication breakdowns between the service provider and the service
requester. The plan defines which user groups are addressed, in which way, and
with what type of information, as shown in Example 4-11.

Example 4-11   Communication plan

The Smart Analytics Cloud provider will give 5 days notice for any
upgrades or infrastructure changes in development and test. This will
be communicated in a common general bulletin to all service requesters.

The Smart Analytics Cloud provider will provide system availability to
the services requesters via a common communication process.

Communication of the status of a reported problem will be per the
common process.

Where possible, an explanation will be provided for outages and
interruptions in service. This will be communicated to the service
requester community as a general communication on the Smart Analytics
Cloud provider web page.

Sevice requesters will be notified of any extended outage via the
Systems Status Notification section of the Smart Analytics Cloud

```
variable web site. Service requesters can select to be notified via
e-mail.

Communication on status will occur via e-mail and postings to a common
repository. The Smart Analytics Cloud project manager and the Smart
Analytics Cloud staff will be responsible for communicating to all
service requesters. All documentation will then be stored in the
project control book for history. Service requesters will be asked to
differ with status within 3 working days or concurrence will be
assumed.
```

4.1.12 Providing user training

While migration to a Smart Analytics Cloud involves moving users to new
reporting tools, user acceptance improves when education is provided. The
cloud provider must decide whether to include education in the fee that the
end-user groups are charged or whether to charge separately for education. If it
is charged separately, the cloud provider can offer education at a special price,
offer to make the necessary arrangements, or can make recommendations about
which training to attend.

4.2 IBM offering for implementing a Smart Analytics Cloud

The IBM Smart Analytics Cloud is a complete end-to-end solution for turning the
corporate deployment of business intelligence and analytics into a common
shared private cloud service that is optimized for the large enterprise customer
by focusing on economy of scale and operational efficiency. This solution
includes hardware, software, and IBM services to deliver a comprehensive
enterprise cloud for large enterprises.

Before offering the Smart Analytics Cloud to the market, IBM developed and
delivered Blue Insight, which is a large scale private cloud deployment of
business intelligence and analytics services. Blue Insight services business
intelligence to 200,000 global IBMers at 56% less costs. Over 1PB of data is fed
through Blue Insight and turned into business insights that allow IBMers to do
their jobs effectively and enables IBM to stay ahead of competition. Blue Insight
serves as the template for the Smart Analytics Cloud.

This section covers:

► A detailed perspective on the Smart Analytics Cloud value proposition

- A view into the components that are offered in the Smart Analytics Cloud
- The importance of IBM services in rapidly achieving the benefits that are offered by the Smart Analytics Cloud

4.2.1 The value proposition of the Smart Analytics Cloud

Businesses are attracted to cloud computing because of the rapid time to value and low costs that are needed to get started with public cloud offerings. However, many public clouds do not afford enterprises a comfortable level of risk. Customers have concerns with the privacy and security that is available in most public cloud offerings and with the lack of operational guarantees, such as service level agreements, which are minimal in many cloud offerings. Because of these concerns, many large enterprise businesses are turning to private clouds to capture the operational efficiency of cloud computing without assuming undue risk.

Private clouds allow enterprises to retain the control they desire and effectively manage the risks that are associated with cloud more easily. Because of the data-centric nature of business intelligence and analytics, the Smart Analytics Cloud is first and foremost a private cloud solution. While services from this solution can be extended and made available to partners, customers, and others, at its core, the design point is to cater to the needs of the large enterprise in a private cloud setting. Large enterprise needs can be specialized when considering the adoption of cloud computing solutions, so the Smart Analytics Cloud speeds the migration of enterprise class workloads, such as, BI and analytics, to a cloud delivery at a rate and pace that is least disruptive to the enterprise.

The value proposition of this offering is geared to solve common problems that large enterprise BI deployments face, such as:

- Adopting cloud computing at a comfortable rate and pace
- Security and privacy of information and sensitive data in the cloud era
- Standardizing business intelligence and analytics workloads
- Delivering intelligence and insight to the users who need it, when they need it
- Maintaining operational budgets while achieving *do more with less* goals

Rate and pace of adoption of cloud services

Many large enterprise companies are introducing cloud computing services into the enterprise. The tendency is to focus on commodity services, such as freely available test resources or other commodity-based cloud offerings. The focus on commodity services has as much to do with the assumed maturity of cloud computing as it does with the ability of the enterprise to support the business trappings of cloud services.

What IBM learned in deploying Blue Insight is that functionally we can build and support the technical implications of delivering business intelligence as a service offering, which was the easy part. What IBM improved on was adjusting the organization(s) to support the pay-as-you-go model that cloud computing proposes.

IBM, like most other large enterprises, is funded using a top-down model. Money starts at the top of the organization and is funneled into the various departments and functional units. There is a rigorous planning cycle for funding that usually involves using historical spending patterns to determine how much money each organization or functional unit will receive for the following year. If the funding for the year (or past years) is not spent, it is lost and likely results in less funding in future years. In this kind of a funding model, variability is not only frowned upon, but it is almost impossible to support.

The Smart Analytics Cloud takes into consideration that cloud computing is not an all or nothing proposal for the large enterprise and that it can be phased, for example, the Smart Analytics Cloud supports a broad range of billing scenarios for the enterprise that starts simply and expands into a more usage-based model, depending on the flexibility of the enterprise. This offering allows our clients to lay a foundation for cloud that is nondisruptive to their business.

Security and privacy of information

Data is foundational to BI. Pulling business insight from data can be the essence of what makes one company successful, while another fails to compete. Data and the impact of cloud computing is a particularly strange beast. The value proposition of the Smart Analytics Cloud, as it pertains to data orientation, is based on three core tenants:

► Private cloud is the key delivery method
► Data is not created equal behind the firewall
► Integration into corporate security is key

Because of the data-oriented nature of business intelligence and analytics, the Smart Analytics Cloud is designed to sit behind your firewall, ensuring that your data never leaves your enterprise. However, data in a large enterprise is often *shades of gray* in terms of ownership and flexibility. Just because data is common to a large enterprise does not mean that there are not many organizations, and each can have a slightly different approach to data management within the large enterprise. Similarly, a large enterprise might not trust the maturity of the public cloud to securely protect their data. The same large enterprise organization might not trust other departments or divisions within the company to manage and run their data.

Another valuable lesson IBM learned in deploying Blue Insight was that organizations want to retain control and ownership of their own data, although

they cannot necessarily harvest its value efficiently. Data is foundational to the organization, and organizations want to own and manage their own data. A single large enterprise can consist of many organizations. While it is organizationally acceptable, albeit challenging, to deliver business insight as a service in the enterprise, this is not the case for data. Organizations prefer to own, manage, and secure their data within the larger enterprise, resulting in silos of data across the enterprise. Traditionally, departmental business intelligence solutions are deployed in or around those data silos, which results in the delivery of narrow business insights that are difficult to aggregate and to deliver comprehensively.

The Smart Analytics Cloud focuses on tapping into the value chain. With a narrow focus on just the BI layer, companies can tap into delivering value and lowering costs, thereby adopting cloud services at a pace that is optimal for the enterprise. Over time, the scope of services that the Smart Analytics Cloud delivers can be expanded as the organization builds trust and value from cloud.

This offering is also designed to take advantage of existing investments, for example, a common investment that most enterprises make is in corporate security, and the Smart Analytics Cloud is designed to integrate into rather than replace those existing investments. Blue Insight taps into the internal group security standards that are in place across IBM, which means that rather than add another layer of security complexity, the Smart Analytics Cloud snaps into the existing investment in the corporation.

Standardizing Blue Insight services

Most large enterprises deploy departmental BI solutions, which are created by grouping business intelligence with data in and around the enterprise, resulting in a multitude of BI deployments. BI solutions tend to be large scale with varying layers of complexity, making the success of each deployment highly dependent on the skills that are assigned to each one. A large enterprise can potentially have ten or more large BI deployments with only one or two flourishing, while the others achieve lesser levels of success.

Standardizing BI is a costly proposition because it involves coordinating many moving pieces, and this is particularly true when the company must also contend with multiple departmental BI deployments.

Cloud computing is based on core tenants, which enable the delivery of services that are massively efficient. A key way that this is possible is through standardizing Web-delivered services. The Smart Analytics Cloud harnesses efficiency by eliminating departmental BI solutions and instead offering a centralized service that is a standard delivery across the enterprise and can be centrally managed and maintained.

Self service

Another key tenant of cloud computing is the ability of consumers to serve themselves. Large enterprise consumers in a private cloud are not like consumers in a public cloud. The end users in a large enterprise do not have money or the ability to pay for a service.

The Smart Analytics Cloud is designed to deal with the unique considerations of large enterprise consumers by providing a customized boarding process. The boarding application supports the ability to board organizations of users, numbering up to a range of 75,000 users, capturing the characteristics of the boarding organizations in an automated fashion, and speeding the boarding time immensely.

With Blue Insight, the time to board an organization of users went from upwards of six weeks to a little over a week. What was once a manual interview process is now an automated process that moves at the rate and pace of the adopting organization. The automated boarding process captures the characteristics of the boarding organization, including information such as:

► How many uses are in the organization?

► Who are the users in the organization?

► What skills does the organization have to support a successful experience with the service, for example report authors?

► Who are the key organization contacts for processes, such as billing, requirements, report authoring, and so on?

After an organization is affiliated with the cloud, the end users of the organization can freely access the cloud by logging in with their username and password.

Doing more with less

The Smart Analytics Cloud is operationally efficient, using less hardware, software, and resources to reach a large audience. Unlike departmental BI, which requires installation costs plus ongoing care and feeding, a Smart Analytics Cloud is a single-centralized deployment that services the entire install base.

In IBM, the savings from Blue Insight were significant over previously installed departmental BI, cutting the costs per user by 56%. The hardware and software savings by centralizing on a single z10™ were greater than $7.75M, while the savings associated with a competency center totaled over $2.5M. In total, IBM saved more than $20M over a five-year period with Blue Insight.

4.2.2 Smart Analytics Cloud component offerings

The Smart Analytics Cloud is a solution that leverages IBM hardware, software, and services. It is the services that make the offering a cohesive and seamless deliverable, so let us start our focus with IBM services.

IBM provides approximately 800 hours of services with the Smart Analytics Cloud to ensure that a fully-functional cloud is available to the client when we leave. We also ensure that the customer can continue successfully with the cloud. Our services are broken into four distinct phases:

► Phase 1: Create awareness of, a strategy for and a governance foundation for BI across the organization

► Phase 2: Prepare for the Smart Analytics Cloud

► Phase 3: Install the base cloud, integrate into the corporate enterprise, and test the cloud use cases

► Phase 4: Optimize the Smart Analytics Cloud

The Smart Analytics Cloud can include incremental hardware upgrades for existing System z customers with capacity, a new System z for customers who are new to the platform, or existing customers who want a new system. The hardware and base-operating environment, including Linux, are aggressively priced and packed as part of the Solution Edition for Enterprise Linux. The hardware components that are included in this offering are:

► A new System z10® or an upgrade to an existing system: IFL with 16 GB memory

► 3-5 years hardware maintenance

► z/VM 5.4 and all base features

► Enterprise Linux

► Connectivity, including 3 4-port FICON® and 2 4-port OSA

The core BI software that is included in this offering is Cognos 8 BI, including the supporting middleware to support the installation, such as WebSphere and DB2.

Software customization packages are available to the clients, depending upon their needs and desires. The identified custom packages are:

► Tivoli monitoring package, including:
  – ITCAM for WebSphere
  – ITCAM for Applications
  – OMEGAMON® XE for z/VM

- Tivoli usage-based billing package:
  - IBM Tivoli Usage & Accounting Manager (ITUAM) for System z
- Tivoli Security package:
  - IBM Tivoli Directory Server

4.2.3 IBM services speeds up benefits from the Smart Analytics Cloud

The Smart Analytics Cloud is first and foremost a services offering with four distinct phases of services. These services help to speed the time-to-value for customers and to ensure that our experiences with Blue Insight are not lost on our customers.

For the first phase, the goal is to work within the client account and create a sustainable business intelligence and analytics strategy and infrastructure plan through the formation of an enterprise-wide competency center for business intelligence and analytics. In this phase, IBM services perform:

- A Cognos technical readiness workshop, which is a two-day workshop to kick off the process, prepare the Cognos architectural landscape, and perform capacity planning
- A Systems Technology Group (STG) technical readiness workshop, which is a one-day architectural planning session
- A Business Intelligence Competency Center (BICC) readiness workshop, which is a four-day roundtable workshop to determine the strategic, technical, and cultural readiness for the BICC journey and to draft a charter, scope, and BICC success plan

In this phase, IBM also provides project management for a consistent and trusted advisor throughout the process.

The second phase is focused on preparing the enterprise for the cloud, which requires deep insight into the existing corporate standards and an understanding of the value chain to create a customized boarding application for the cloud. In this phase IBM delivers:

- Corporate standards assessment: Discovery sessions that are focused on capturing the existing corporate standards around security and automation in particular
- Corporate value assessment: Discovery sessions that are focused on capturing the existing corporate organization and pertinent roles to the cloud, and a view into performance stakeholders in the value chain

- ▶ Optional migration services: Plan the migration of existing enterprise business intelligence reports, dashboards, and other objects from vendor products

In the third phase, IBM services actually gets the cloud up and running. Throughout the process IBM delivers:

- ▶ Installation and configuration of the System z foundation
- ▶ Installation and configuration of the existing business intelligence and analytics environment
- ▶ Creation and installation of the service planning and user boarding application, which is customized to the corporation
- ▶ Custom enterprise integration with alignment and technical integration into the existing corporate enterprise standards for security, automation, and other IT enterprise-wide initiatives
- ▶ Custom component implementation, such as the installation and configuration of optional Tivoli components, such as monitoring, metering, and security

In the fourth and final phase, IBM services focuses on optimizing the Smart Analytics Cloud for the enterprise. In this phase, our services team delivers a BI-operational administration knowledge transfer that is designed to provide support, direction, and documentation for customers who want to maintain their own cloud. The goal is to make customers so familiar with the product that they can, for the most part, independently manage their system.

4.2.4 Conclusion

The Smart Analytics Cloud can deliver significant value to large enterprise customers who want to harness the value of cloud computing at a comfortable rate and pace. Using our experiences with Blue Insight, which serve as a template for the Smart Analytics Cloud, IBM now offers this solution to you. Our own $20M cost reduction can be yours.

Part 3

Architecture

In this part, we describe the architecture for the Smart Analytics Cloud:

► We give an architecture overview and describe how the cloud fits into a larger enterprise context.

► We drill down in Chapter 6, "Functional architecture" on page 63, into a functional architecture of the cloud and the rationale behind using it. We describe the functions that are required not only for the IBM Cognos 8 BI components, but also for the onboarding functionality that is used as an approach self-service access to IT functionality.

► In Chapter 7, "Operational architecture" on page 73, we give an operational architecture overview of the cloud that shows how the required functions are placed on actual infrastructure and which technology is chosen.

► We look into the systems management aspects, which ensure the vitality and the automation of the cloud.

5

Architecture overview

In Part 2, "Business" on page 17, we introduced the business objectives of the Smart Analytics Cloud. In this part, we discuss the technical architecture that is required to build a Smart Analytics Cloud.

The Smart Analytics Cloud operates in a data warehousing environment. To show how it fits into a medium-to-large-sized enterprise, we look into the next larger architectural context in 5.1, "Data warehouse environment" on page 48.

Narrowing down the focus from the overall data warehouse to the Smart Analytics Cloud, we develop a system context in section 5.2, "System context of the Smart Analytics Cloud" on page 50.

Based on the data warehouse architecture and the system context, we drill down in section 5.3, "Functional overview" on page 53 into a functional architectural overview of the cloud.

In section 5.4, "Operational overview" on page 59, we discuss some alternatives to deploying on Linux on System z to give an operational architecture overview of a typical Smart Analytics Cloud implementation.

5.1 Data warehouse environment

While data warehousing is an optional element in the overall Smart Analytics Cloud architecture, here we include a discussion for those of you who already included or are thinking of including one in your cloud.

Over the past 15 to 20 years data warehousing and business intelligence evolved from just retrieving data from transactional systems and compiling some reports into a highly sophisticated data warehouse provisioning process. It transforms the extracted data into an enterprise asset and turns it into valuable information.

The supply process for data warehouses has become more and more understood and standardized. However, the reporting part is still handled on a department level. More harmonization and standardization concerning the architecture, the tools, and the interpretation of the retrieved information is needed.

5.1.1 Data warehouse architecture

To put the Smart Analytics Cloud into a larger context, we first look at a data warehouse environment. Figure 5-1 on page 49 shows a common enterprise data warehouse architecture consisting of nine layers. This architecture follows a hub-and-spoke architectural style.

There are a number of different flavors of data warehouse architectures, dealing with operational data stores, direct access to the data warehouse and virtual data warehouses, which we do not cover here because it exceeds the scope of this book. For the rest of this book, we follow the definitions from *Enterprise Data Warehousing with DB2 9 for z/OS*, SG24-7637 and references mentioned in that book.
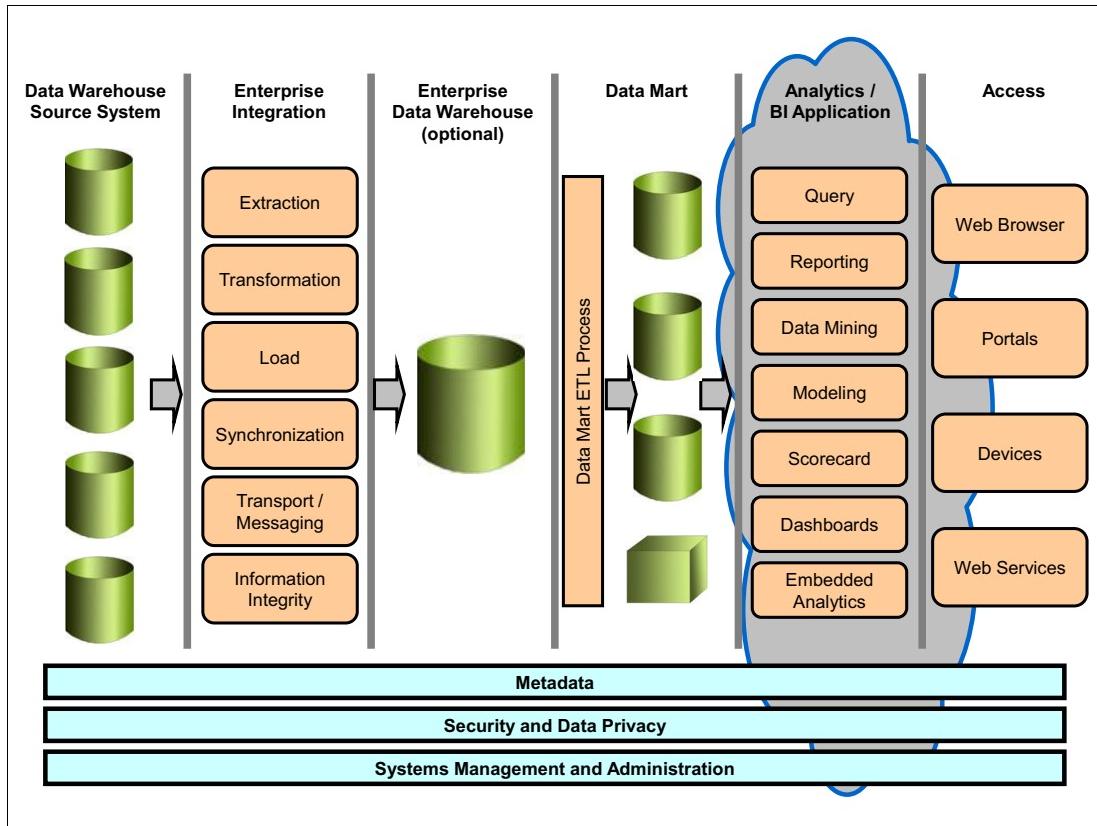
Figure 5-1   Data Warehouse Architecture overview

Description of the layers:

► The *Data Warehouse Source System* layer is composed of all source systems for the data warehouse. These systems are characterized as having Online Transaction Processing (OLTP).

► The *Enterprise Integration* layer is responsible for extracting and processing the data from the source systems and performing the data warehouse provisioning.

► The *Enterprise Data Warehouse* layer consists of the centralized, corporate-wide consolidated data warehouse as the data hub.

► A *Data Mart* is a subject oriented data store that provides an application centric optimized view of the warehouse data. Typically, a data mart has an Online Analytical Processing (OLAP) system and data is stored in a normal relational database.

Populating a Data Mart is much the same as populating a data warehouse, but as the data warehouse has already made up a consolidated and integrated enterprise view of the data, this process is much easier.

► The *Analytics Business Intelligence Application* layer organizes all applications that operate on the data of the data mart layer. It is not uncommon for several applications to share the same data mart. Sample business intelligence applications include:

  – Campaign management
  – Personalization and adaptive marketing
  – Corporate dashboards
  – Customer and market segmentation and scoring
  – Channel effectiveness
  – Warranty analysis
  – Risk management
  – Operations and product analysis
  – Cost analysis
  – Compliance reporting
  – Fraud detection

> **Note:** As shown in Figure 5-1 on page 49, the Smart Analytics Cloud deals mainly with the Analytics Business Intelligence Application layer.

► The *Access* Layer links access methods with the Business Intelligence Applications and Analytical Methods. Access methods, such as web-based front-ends, rich clients, or system interfaces, incorporate the results of a business intelligence application or an analytical function into an OLTP system.

To round out the architecture, we need some cross functionality, such as metadata management, security and data privacy, and several systems management disciplines.

5.2 System context of the Smart Analytics Cloud

In this section, we review the interfaces of the Smart Analytics Cloud using external systems called *System Actors* (silicon-based) and human users called *Human Actors* (carbon-based). We attach actors to typical non-functional requirements.

Our starting point is the *Analytics / BI Application* layer of the DWH reference architecture, as shown in Figure 5-1 on page 49. When selecting this and the

nearby layers and doing a little refinement, we get the system context, as shown in Figure 5-2.
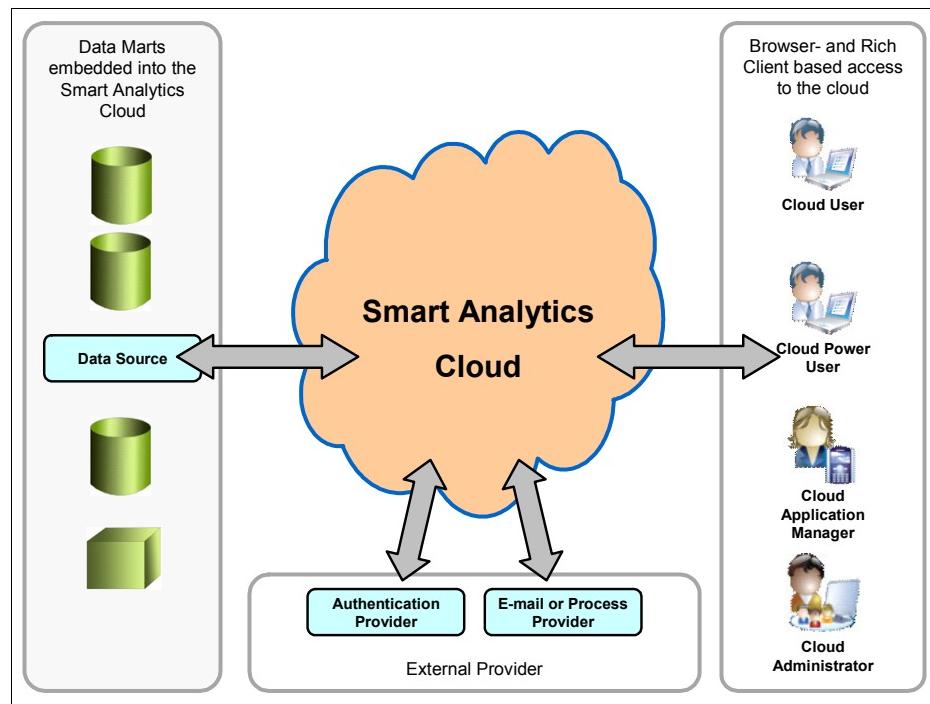
Figure 5-2   Smart Analytics Cloud system context

Actor: Cloud User

Each user who wants to use the Smart Analytics Cloud uses his browser to access the reports.

His non-functional requirements depend on the non-functional requirements of the cloud-application he wants to use and of the non-functional requirements the underlying data sources provide. In 4.1, "Scope of the Smart Analytics Cloud" on page 28, we gave you a sample service definition.

Typical requirements:

► Availability: Office hours for each user
► Quantity: Thousands, possibly every user in your enterprise

Actor: Cloud Power User

Power Users need the Cognos Framework Manager GUI application to access the Framework Manager, Metrics Manager and Transformer of Cognos.

His non-functional requirements depend on the non-functional requirements of the cloud-application he wants to use and of the non-functional requirements the underlying data sources provide. In 4.1, "Scope of the Smart Analytics Cloud" on page 28 we have given you a sample service definition.

Typical requirements:

► Availability: Office hours for each user
► Quantity: A few

Actor: Cloud Application Manager

Because new Business Intelligence applications can join the cloud, a user must register the application and manage the onboarding process. In this case, the user acts as a *Cloud Application Manager*. This actor uses a browser to access the cloud.

In general, this actor's non-functional requirements (especially the availability requirements) are less challenging as, for example, the reporting user.

Typical requirements:

► Availability: Office hours for each user
► Quantity: Few

Actor: Cloud Administrator

This is a generic actor for all administrative tasks needed in the solution, including, but not limited to:

► Administration of the onboarding process
► Application server administration
► Administration of the reporting components
► Database server administration
► System Management tasks like security, monitoring, and so on

This actor's non-functional requirements are based on normal working hours.

Typical requirements:

► Availability: Office hours for each user
► Quantity: A few

Actor: Data Source

This system actor is in fact a placeholder for every data source you will need for your applications in the cloud, which includes:

► Data marts based on relational databases
► Data marts based on cubes (OLAP, ROLAP)

- ► Data warehouse
- ► Data of reference systems
- ► External data

These systems are out of the management scope of the cloud because they are normally individually owned by the user groups that are using the cloud.

Typical requirements:

- ► Availability: Individually defined by each system
- ► Quantity: A few

Actor: Authentication Provider

Your cloud will probably use an existing authentication provider infrastructure that is represented by this system.

Normally the authentication service has high non-functional requirement characteristics.

Typical requirements:

- ► Availability: Individually defined by each system

- ► Quantity: One or two, depending how you have structured your authentication infrastructure

Actor: Email Provider

This system actor is an existing email system or business process engine. The Smart Analytics Cloud uses this actor to send e-mails or to interact with an existing process engine as, for example, within the onboarding process to inform participants about status changes.

Normally an email provider has high non-functional requirements characteristics.

Typical requirements:

- ► Availability: Individually defined by each system

- ► Quantity: One or two, depending how you have structured your email infrastructure

5.3 Functional overview

The IBM offering for the Smart Analytics Cloud is based on IBM Cognos 8 BI on System z to deliver a cloud infrastructure. The offering focuses as an initial step to the Analytics and BI Application layer of the overall data warehouse

architecture. Three Cognos capabilities from the overall Cognos architecture shown in Figure 5-3 are selected and are implemented in our environment:

► Query
► Reporting
► Analysis

Further capabilities can be included in later releases of the Smart Analytics Cloud or implemented outside the cloud as a separate offering.
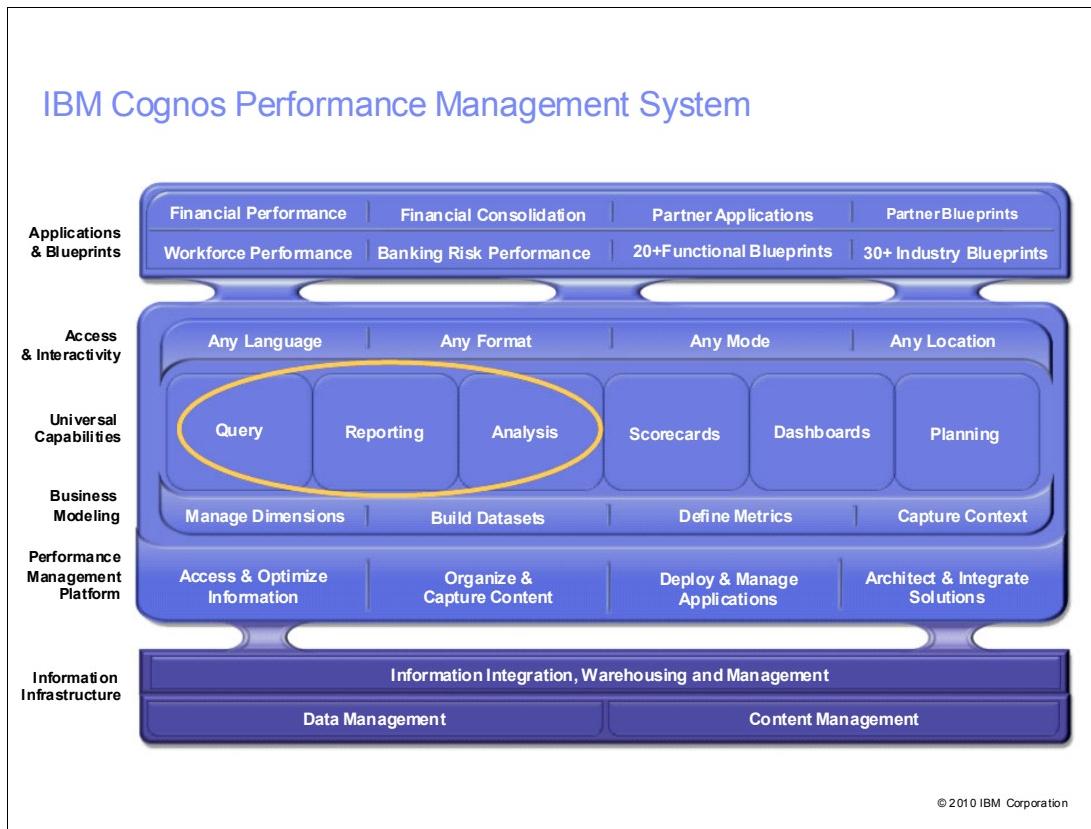
Figure 5-3   IBM Cognos 8 BI architecture overview

5.3.1 Smart Analytics Cloud

To provide the selected Cognos elements in a cloud, components for provisioning, self service, and monitoring are required.

When combining the Cognos functionality and the required components to set up, operate, and manage the cloud, we get an architecture overview for a Smart Analytics Cloud.

In addition to the Analytics Cloud functions and the Cloud Management Functions, a reliable, available infrastructure is needed to support the accelerated provisioning process. These elements are shown in Figure 5-4.
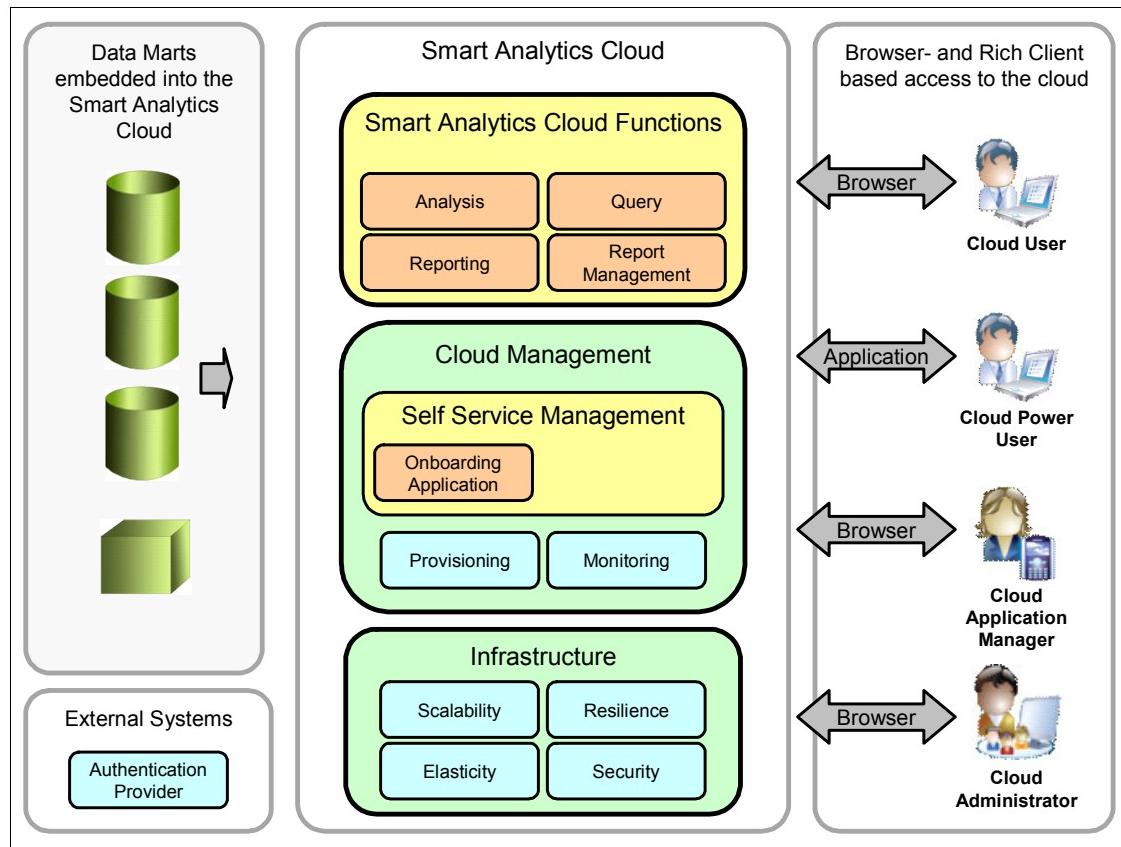
Figure 5-4   Smart Analytics Cloud Architecture Overview

The Smart Analytics Cloud consists of the following building blocks (as described in section 5.3.2, "Description of the building blocks" on page 56):

► Smart Analytics Cloud Functions
► Cloud Management:
    – Self Service Management
► Infrastructure

The data marts on the left side in Figure 5-4 on page 55 are not part of the Smart Analytics Cloud because their structure is individual for each user group. Therefore, they are not suitable to be provided in a standardized cloud. Authentication is also provided externally because existing corporate-wide systems can easily be used.

The right side of Figure 5-4 on page 55 shows the clients and applications that use the Cognos elements. The type of client depends on the type of task that has to be accomplished. Because this is individualized for each user group, these components are currently not provided in a cloud.

5.3.2 Description of the building blocks

In this section, we describe the building blocks.

Smart Analytics Cloud functions

This is the functional core of the Smart Analytics Cloud that stores and executes the functionality of all applications who enter the cloud. The main functional and non-functional requirements of this block are described in section 4.1, "Scope of the Smart Analytics Cloud" on page 28.

In our example in this book, all functionality is implemented by IBM Cognos 8 Business Intelligence product.

Cloud Management

Essential for every application on an enterprise scale are system management functions, which we named *Cloud Management*.

We emphasized *provisioning* because this is essential for the onboarding of new applications and users. Increasing the workload by adding users and applications must be planed and implemented by a highly automated process.

Self-service management

Because new BI applications can join the cloud and new users can register themselves to access the cloud, we have to provide an *Onboarding Application* that manages these processes and interacts with an existing authentication provider based on RACF®, Tivoli Directory Server, or a comparable element to retrieve user information and access rights of each user, which will be mapped onto roles within the cloud.

This building block belongs to the overall Cloud Management.

Infrastructure

From an infrastructure perspective, cloud computing represents an evolutionary shift that, while incorporating some new ideas, also ties together many existing concepts to solve today's challenges. Because its core strengths best meet the cloud infrastructure requirements, System z is the platform used for the Smart Analytics Cloud. Key aspects of a private cloud infrastructure are *scalability*, *elasticity*, *security*, and *resiliency*:

Scalability
Part of the challenge for cloud computing is creating an infrastructure that appears to provide infinite compute resource. Scalability is the ability for a platform to grow to meet user demand. System z has the ability to non-disruptively add resources and scale up to handle significant capacity within a single server.

Elasticity
In a cloud environment, elasticity is perhaps even more important than scalability. Elasticity represents the ability to adjust capacity up or down without disrupting business processes. System z's Capacity on Demand functionality has helped user to adjust for seasonal peak workloads and has an established cost structure that clients are comfortable with. Elastic compute capacity, coupled with sound economic structure, is a core enabler of the cloud computing.

Security
In a private cloud, it is critical to have a secure virtual infrastructure. System z has the highest security classification for a general purpose server. The platform's LPAR or logical partition technology meets the EAL5 certification, qualifying that each logical partition keeps data secure and distinct from others.

Resiliency
While some platforms were designed originally for academic or other purposes, System z was originally designed for business. Over the years, redundant components and resiliency technologies have evolved and become deeply embedded into System z design. System z is the premier platform in ensuring business systems stay up through disaster recovery, repair and upgrade, and also software and application changes.

5.3.3 Data Sources

This layer maps to the *Data Mart* Layer of the *Data Warehouse Architecture Overview.* All Source Data for the Business Intelligence Applications and Analytical Functions needed are referenced here and thus become part of the cloud.

Access to these systems is done by well established techniques, for example (but not limited to):

► DB2 to access data residing in DB2 on z/OS®, Linux, AIX®, Windows® and various Unix-derivates

► JDBC access for non-DB2 database management systems

► IBM Cognos Virtual View Manager uses standard JDBC drivers, but supports more data sources on Linux on System z

► IBM InfoSphere™ Federation Server or IBM InfoSphere Classic Federation for access to non-relational data and non-DB2 database management systems

For the Smart Analytics Cloud it is not necessary to distinguish whether a source systems is a data mart, the enterprise data warehouse, an operational data store or any other database management system.

For a detailed and updated list of data sources that can be accessed by way of IBM Cognos Virtual View Manager, refer to:

http://www.ibm.com/support/docview.wss?rs=3442&uid=swg27014427

For a detailed and updated list of data sources that can be accessed by way of IBM InfoSphere Federation Server, refer to:

http://www.ibm.com/support/docview.wss?uid=swg27015299

For a detailed and updated list of data sources that can be accessed by way of IBM InfoSphere Classic Federation Server, refer to:

http://www.ibm.com/support/docview.wss?uid=swg27011950

5.3.4 Cloud access

The Smart Analytics Cloud leverages the facilities of IBM Cognos 8 BI of lightweight access using web-based technology for *Cloud Users*. Only *Cloud Power Users* who work with source data model metadata need a rich client with IBM Cognos 8 Framework Manager.

All human actors are described in detail in section 5.2, "System context of the Smart Analytics Cloud" on page 50.

5.4 Operational overview

First we consider some general aspects for developing an operational model. Based on that we present an overall operational overview for a multi environment solution including development, testing, and production.

For details of the operational architecture and a discussions how to fulfill non-functional requirements, review Chapter 7, "Operational architecture" on page 73.

5.4.1 General considerations for an operational model

The underlying platform for the Smart Analytics Cloud is System z. The core strengths of the platform—virtualization, scalability, elasticity, security, and resiliency—provide a robust infrastructure from which to develop an enterprise class private cloud. Cognos 8 BI and other software components can be deployed on one or more System z systems and can use System z's virtual infrastructure capabilities.

The platform's virtualization infrastructure is built upon two technologies, z/VM and LPAR:

► z/VM is a software-based hypervisor that allows a single System z to run hundreds, or potentially thousands, of virtual servers. z/VM has an eloquent paging subsystem that commonly results in virtual-to-real memory ratios of 2.5-to-1. Sharing memory among applications results in better resource utilization and further reductions in costs. z/VM also incorporates a high-throughput virtual network that, when applied to server consolidation, can reduce requirements on physical hardware such as switches and routers.

► LPARs or logical partitions are secure subsystems within a System z. Using hardware mechanisms, System z divides the hardware on one server one or more LPARs. All software runs in a LPAR and the partitioning mechanisms have negligible performance impact. It is important to differentiate a LPAR from a hypervisor like z/VM in that the LPAR represents partitions of hardware rather than a layer of software that virtualizes it. Processor, I/O, and network resources can be shared or dedicated to LPARs. Each LPAR is assigned its own memory. Direct communication across LPARs can be done using HiperSockets™, an in-memory network for virtual servers within one system.

Deployment

In a typical deployment, all applications are deployed as guests within a single z/VM. While it is possible to run applications directly in a LPAR, running as guests under z/VM is the preferred approach. Leveraging z/VM to oversubscribe processor resources, even though z/VM incurs some overhead, can result in more effective utilization and improve overall application throughput. z/VM also provides the flexibility to easily deploy new instances of servers quickly and on demand. Figure 5-5 shows a typical Cognos 8 BI server deployment using z/VM.
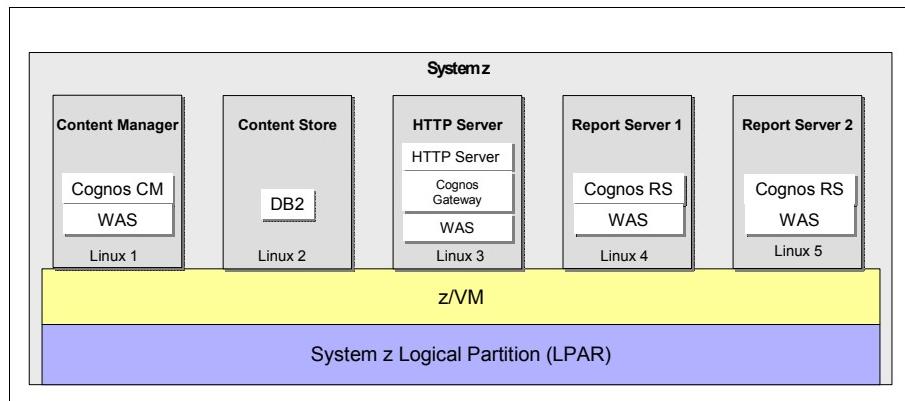
Figure 5-5   Typical Cognos 8 BI server deployment using z/VM and multiple Linux guests

A System z-based private cloud infrastructure offers flexibility beyond allowing multiple Linux operating systems to run within a single LPAR. You might choose to set up multiple LPARs, perhaps for increased high availability or for more control over resource allocation among environments, for example, you might set up one LPAR for production and another for development and test. You can also have two production LPARs and deploy components in each, using load-balancing routers to provide increased availability between two LPARs.

> **Note:** Additional information about high-availability architecture on System z is in *Achieving High Availability on Linux for System z with Linux-HA Release 2*, SG24-77111.

System z virtual infrastructure

Figure 5-6 on page 61 depicts a model of the System z virtual infrastructure and shows how resources, including general purpose processor cores (GPs) and Integrated Facility (IFL) for Linux, can be virtualized and shared. An IFL is a processor core that is dedicated to running Linux.
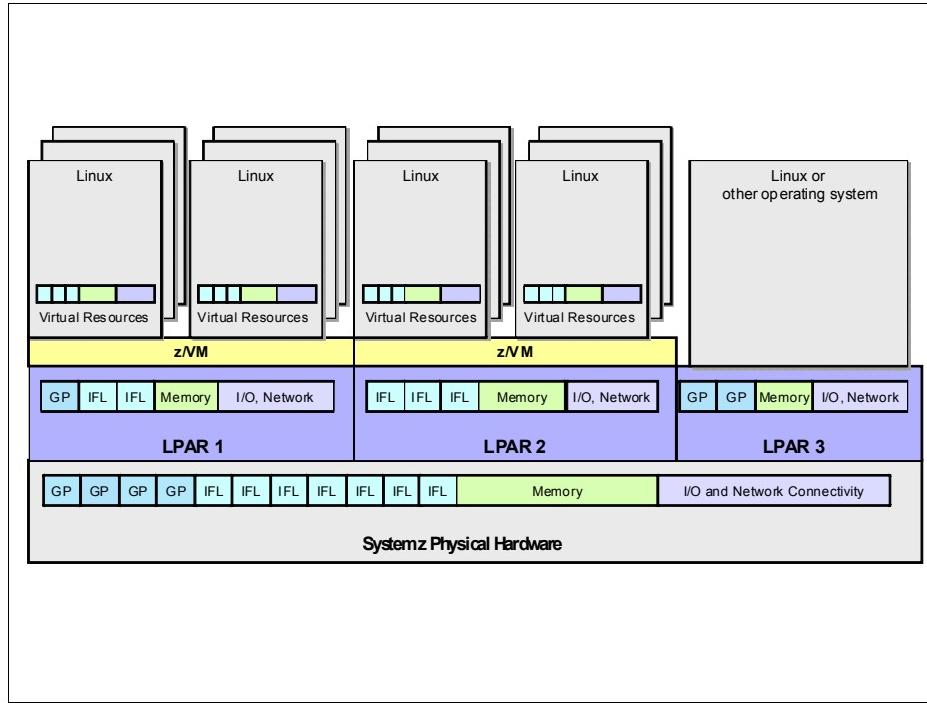
Figure 5-6   System z virtualized infrastructure

5.4.2 Operational overview for multiple deployment environments

You must separate the development, test, and production stages into different environments because they will have different sets of non-functional requirements and require different implementations and configurations, as shown in Figure 5-7 on page 62.

Depending on your requirements, you might come to different solutions, for example, some clients have different networks for production, development, and test.
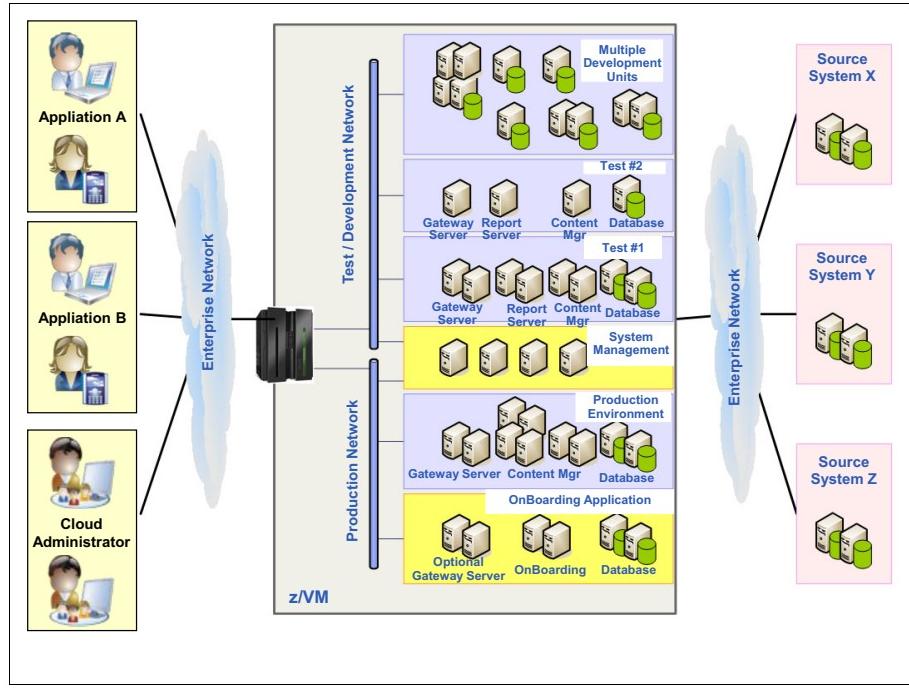
Figure 5-7   Operational model for multiple environments

In this example, we have multiple applications with their users (*Cloud User* or *Cloud Power User*) and the users acting in the role as *Cloud Application Management*. But for the cloud there is only one user group that will act in the different roles of Cloud Management.

All applications have their own development environment. For test, we have two environments (performance and functional test) and one production environment for all applications.

The whole cloud has one onboarding application and all environments share the same systems management functions.

In section 7.2.4, "Multiple deployment environments" on page 83 we take a deeper look into the requirements of these environments and discuss some alternatives.

6

Functional architecture

In this chapter, we take, from a functional point of view, a closer look into the building blocks of the Smart Analytics Cloud, which were identified in Chapter 5, "Architecture overview" on page 47. We break up the building blocks into components whose functionality is deployed in Chapter 7, "Operational architecture" on page 73 onto nodes of the operational model:

► We give a rough overview of the *Smart Analytics Cloud functions*. Although we will implement it using IBM Cognos 8, we need additional components to work within the operational model.

► The core part of *Self-service management* is the *Onboarding Application*, which interacts with the provisioning of *Infrastructure.* We describe the onboarding application in this chapter in more detail.

All other functionality of the building blocks *Infrastructure* and *Cloud Management* is discussed in more detail in Chapter 8, "Cloud Management architecture" on page 97.

6.1 Smart Analytics Cloud functions

The application building block, *Smart Analytics Cloud functions*, is the functional core of the Smart Analytics Cloud and is implemented using IBM Cognos 8 Business Intelligence. Because IBM Cognos 8 BI is a packaged solution, we do not need to re-model the components in the IT Architecture. However, to make placement decisions in the operational model, we must introduce the components that are later used. Figure 6-1 shows an overview of the Cognos software components.
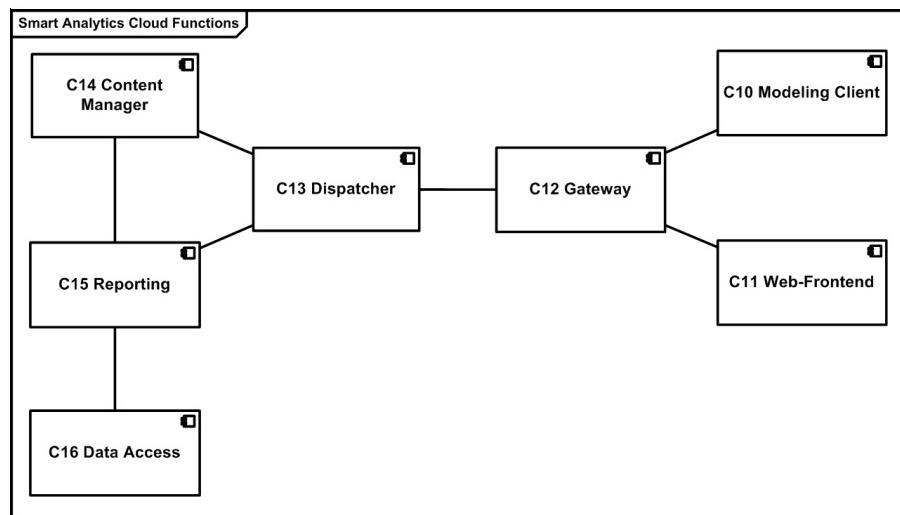
Figure 6-1   Component model smart analytics cloud functions

6.1.1 Component descriptions

In this section, we describe the components in Figure 6-1.

C10 Modeling Client
This component is implemented by the IBM Cognos 8 Framework Manager. It is needed to model the metadata, which describes the source systems and their underlying data models.

C11 Web-Frontend
This component handles the browser-based web-front end interface for a typical user who wants to run reports and work with data.

C12 Gateway

This component is implemented by the IBM Cognos 8 Gateway. It manages all web communication for the IBM Cognos 8 Platform. The workload on the IBM Cognos 8 Gateway server is comparatively lightweight, therefore, it requires minimal processing resources.

C13 Dispatcher

This component is implemented by the IBM Cognos 8 Dispatcher. It performs the load balancing of requests from the clients. It is a lightweight Java™ servlet that manages (and provides communication between) application services.

C14 Content Manager

This component is implemented by IBM Cognos 8 Content Manager, which manages the storage of customer application data, including security, configuration data, models, metrics, report specifications, and report output. It is needed to publish models, retrieve or store report specifications, handle scheduling information, and manage the Cognos name space. Content Manager maintains information in a relational database that is referred to as the content-store database.

C15 Reporting

This component is implemented by IBM Cognos 8 Report Server. It is responsible for processing the report or query service. It is started dynamically by the *C13 Dispatcher* as needed to handle the request load.

C16 Data Access

This component provides the necessary access mechanisms to the source systems for the *C15 Reporting*. Possible access mechanisms are described in 5.3.3, "Data Sources" on page 58. To keep the component model simple we will not delve deeper into this component.

6.2 Onboarding application

The onboarding application is an essential part of the onboarding process, as discussed in Part 5, "Driving the cloud" on page 247.

It automates and standardizes the process and allows the end-users to self service their request. The self service shortens the provisioning process time from weeks to a couple of days.

> **Note:** We will discuss a generalized version of an onboarding application. Depending on your established processes, you will need a tailored and customized version of this application.

Typical use cases for this application are:

► Register an Organization
► Register an Application
► Manage approval process for an Application
► Deploy additional hard- and software using an automatic provisioning

6.2.1 Component model

As discussed in 3.2, "Cloud computing" on page 21, the concept of cloud computing includes self service mechanisms that allow users to request a service tool. The onboarding application is an approach to this self service concept. It supports a standardized process to apply to onboard on the cloud application. The process used for onboarding is described in Chapter 16, "Onboarding" on page 253.

This section describes the component model of the onboarding application, as depicted in Figure 6-2.
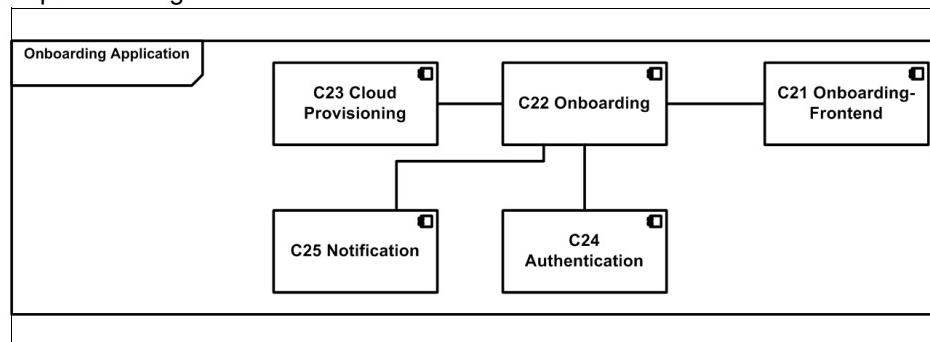
Figure 6-2   Component model onboarding application

C21 Onboarding-Frontend

This component is responsible for presenting the user interface and managing the user input.

C22 Onboarding

This component is the core component of the onboarding application, which manages the data and implements the necessary steps of the process.

C23 Cloud Provisioning

An important step within the onboarding process is the provisioning of new or changed hardware and software configurations. This component is the bridge between the onboarding application and the systems management functionality.

C24 Authentication

This component integrates with an existing authentication provider (an existing LDAP-directory or RACF) to authenticate the user and to retrieve its authorizations.

C25 Notification

The onboarding application integrates into an existing workflow-management system or only an email infrastructure to inform participants about status changes and to trigger the next steps of the process.

6.2.2 Data model

The onboarding application uses a data model as the high-level data model, shown in Figure 6-3 on page 68, to capture the necessary information and to manage the onboarding process. The entities are described in the next section.
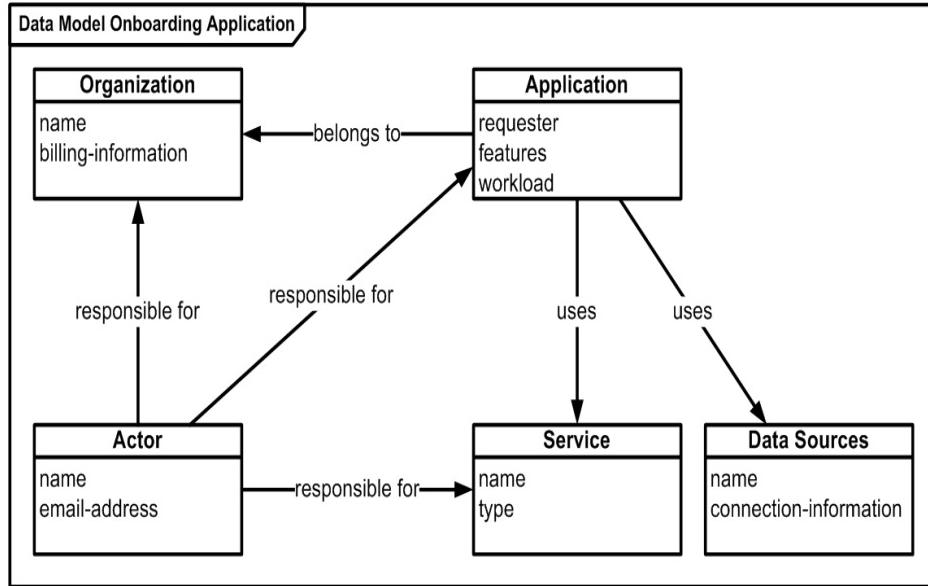
Figure 6-3   Data model onboarding application

Data sources

An application needs to access certain data sources. This entity collects the needed sources, not considering whether they are already connected to the Smart Analytics Cloud or not.

Service

An application can order certain services during the onboarding process from a *Service Catalogue* (see Chapter 16, "Onboarding" on page 253). Each *Service* must have a contact person who is responsible for planing its use.

A Service in the context of the Smart Analytics Cloud can be one or any combination of the following types:

► Onboarding of an *Organization*: Registering a new organization for the Smart Analytics Cloud

► Infrastructure: Providing environments for development, test, and production

► Services: Consulting services to onboard on the cloud

► Implementation: Migrating existing reports in other BI tool versions to Cognos 8 BI report format

Actor

An actor is a human person who performs certain tasks within the onboarding process, as documented in section 16.1, "Roles in the onboarding process" on page 254. Typically you have:

- ► Service Requester
- ► Service Request Approver
- ► Service Manager
- ► Smart Analytics Cloud Administrator
- ► Smart Analytics Cloud Consultant

Organization

An organization a business unit or group within the enterprise that owns applications or sets of reports and wants these to be included in the Smart Analytics Cloud. Typically an Organization has more than one application to submit to the cloud. This entity holds contact (using the actor) and billing information.

Any user can initiate the onboarding or registering of an organization to emphasize the self-service management paradigm instead of preregistering all organizations or making them editable only by the *Smart Analytics Cloud Administrator.* After an organization is registered, an application can reference the organization and the billing information that is contained within.

In Figure 6-4 on page 70, we show an example state diagram for the organization entity. Depending on your onboarding process, the diagram can vary.
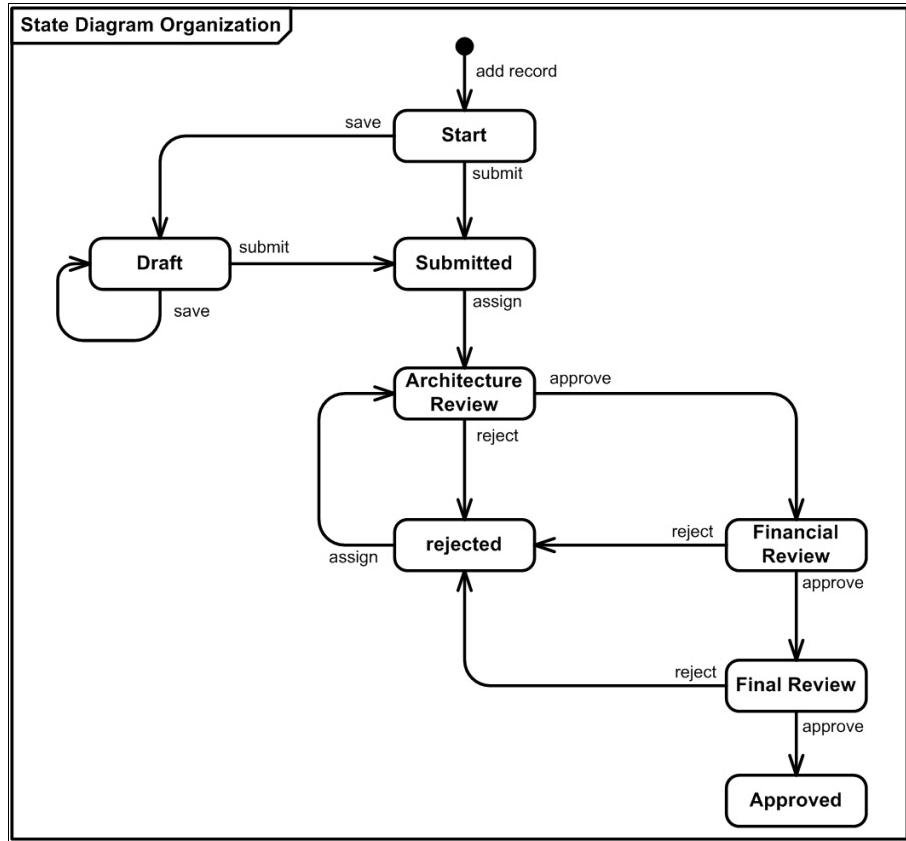
Figure 6-4   State diagram organization

A user adds (status *Start*), edits, and saves (status *Draft*) an organization. After finishing, it is submitted (status *Submitted*) and reviewed by the *Smart Analytics Cloud Administrator* (see section 16.1, "Roles in the onboarding process" on page 254) who evaluates the organization from an architecture and financial point-of-view:

► How many potential users must be added to the environment?
► Are the potential new applications fit for the cloud?
► What are the expectations of the new organization?
► Are the financial details sufficient for billing?

Application

An application consists of a number of reports that access defined source systems. Any user acting in the role of a *Service Requester* can initiate the onboarding of an application.

Typically, the *Service Requester* must collect and document the following information for the approval process:

► A short description of the application, including the intended time frame of development and production start

► Accessed data sources

► Contact information of project manager and architect

► Estimated workload, including the umber of users, number of reports, schedule of report execution, and the size of the result sets

► Needed features and tools of IBM Cognos 8 BI

Figure 6-5 is an example state diagram for an *application* when the *Service Requester* requests three environments for development, test, and production. This diagram is an example of having three cloud environments and needing to keep track of the phases that deal with each environment. Depending on your onboarding process, your state diagram can vary.
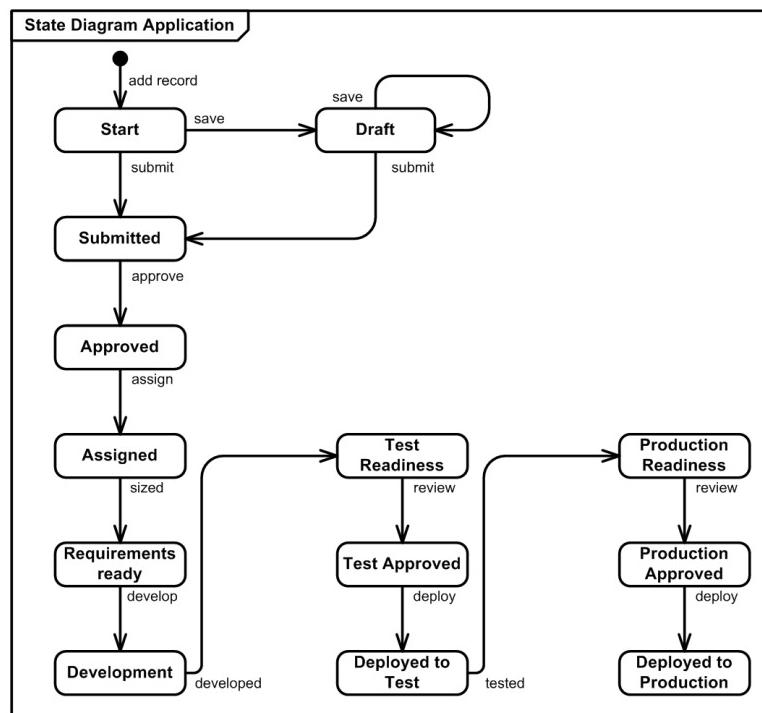
Figure 6-5   State diagram application

The *Service Requester* adds an application (status *Start*), saves the record (status *Draft*), and submits it (status *Submitted*) for approval by the *Service*

Request Approver who must approve the request for further processing (status *Approved*). For a description of the defined process roles, see section 16.1, "Roles in the onboarding process" on page 254.

The application runs through the following status before it is finally deployed into production:

▶ Status *Assigned*: The *Service Manager* is assigned for the application as a focal point and guides the application through the rest of the process.

▶ Status *Requirements ready*: Requirements gathering finished, and the result is documented and filed in the onboarding application. Now the *Service Manager* initiated the set up of the development environment.

▶ Status *Development*: The development cloud is established, and the application can be developed by the service requester.

▶ Status *Test Readiness*: The application is developed and ready for test. Testing can occur after a test readiness review.

▶ Status *Test Approved*: The review has a positive result.

▶ Status *Deployed to Test*: The application can be deployed to test and tested.

▶ Status *Production Readiness*: The application is tested and waits for the production readiness review.

▶ Status *Production Approved*: The review has a positive result.

▶ Status *Deployed to Production*: The application can be deployed into production.

7

Operational architecture

In this chapter, we describe the operational architecture for the Smart Analytics Cloud. The aspects that are relevant to system management are elaborated in more detail in chapter Chapter 8, "Cloud Management architecture" on page 97:

► We place the functional elements of the component model (documented in Chapter 6, "Functional architecture" on page 63) onto logical nodes. We attach to the nodes typical non-functional requirements derived from section 4.1, "Scope of the Smart Analytics Cloud" on page 28.

   This model serves as a reference architecture for your solution.

► Before we give you the operational model for our lab environment, we discuss non-functional requirements and their implications on the architecture and how to cope with them on an architectural level.

► We develop from the logical operation model the physical operational model for our lab environment.

► We walk through the physical operational model for the nodes that are responsible for the core cloud reporting functionality that we described in "Smart Analytics Cloud functions" on page 56.

7.1 Logical operational model

To have greater flexibility in modeling the physical operational model, we separate the components as far as possible and place them on a number of different logical nodes, as show in Table 7-1. Sometimes we must separate the data aspect of a component from its execution aspect because the data can (and will be) stored on a different node. Because of the needed systems management functionality, we introduced additional logical nodes that do not host application functionality.

Table 7-1   Mapping components to logical nodes

Node	Components
LN1 Gateway	C11 Web-Frontend C12 Gateway
LN2 Reporting	C13 Dispatcher C15 Reporting (execution aspect) C16 Data Access
LN3 Content Manager	C13 Dispatcher C14 Content Manager (execution aspect)
LN4 Metadata	C14 Content Manager (data aspect) C15 Reporting (data aspect: reports)
LN9 Modeling Client	C10 Modeling Client
LN11 Deployment Manager	This node is introduced because of WebSphere administration purposes
LN12 SM Presentation	This node is introduced for presentation purposes for systems management functionality
LN13 SM Monitoring	C23 Cloud Provisioning plus additional systems management components
LN21 Onboarding	C21 Onboarding-Frontend C22 Onboarding (execution aspect) C24 Authentication C25 Notification: This component is not used because we do not integrate into an existing email-infrastructure
LN22 Onboarding Data	C22 Onboarding (data aspect)

7.1.1 Overview of the logical operational model

The logical operational model follows the three-tier architecture of Cognos consisting of a data, application, and presentation tier, as documented in the online documentation at:

http://publib.boulder.ibm.com/infocenter/c8bi/v8r4m0/index.jsp

For a quick overview, visit the following site:

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101437

We added an additional tier for the human actors and grouped the nodes into three submodels:

► The *Onboarding Application* contains the nodes and actors necessary to perform the onboarding process.

► The *Reporting* is the core piece of the Smart Analytics Cloud environment that delivers the reporting functionality and access to the data sources that are needed.

► The *System Management* collects all nodes that are needed to administer and monitor the solution. Because these nodes interact with every other node in the architecture, the connections are not drawn.

Depending on the different possibilities you have when you advance from the logical to the physical operational model, you deploy these submodels in separate ways (7.2.4, "Multiple deployment environments" on page 83).

We reused the actors, as defined section 5.2, "System context of the Smart Analytics Cloud" on page 50.

Figure 7-1 on page 76 shows a logical operational model.
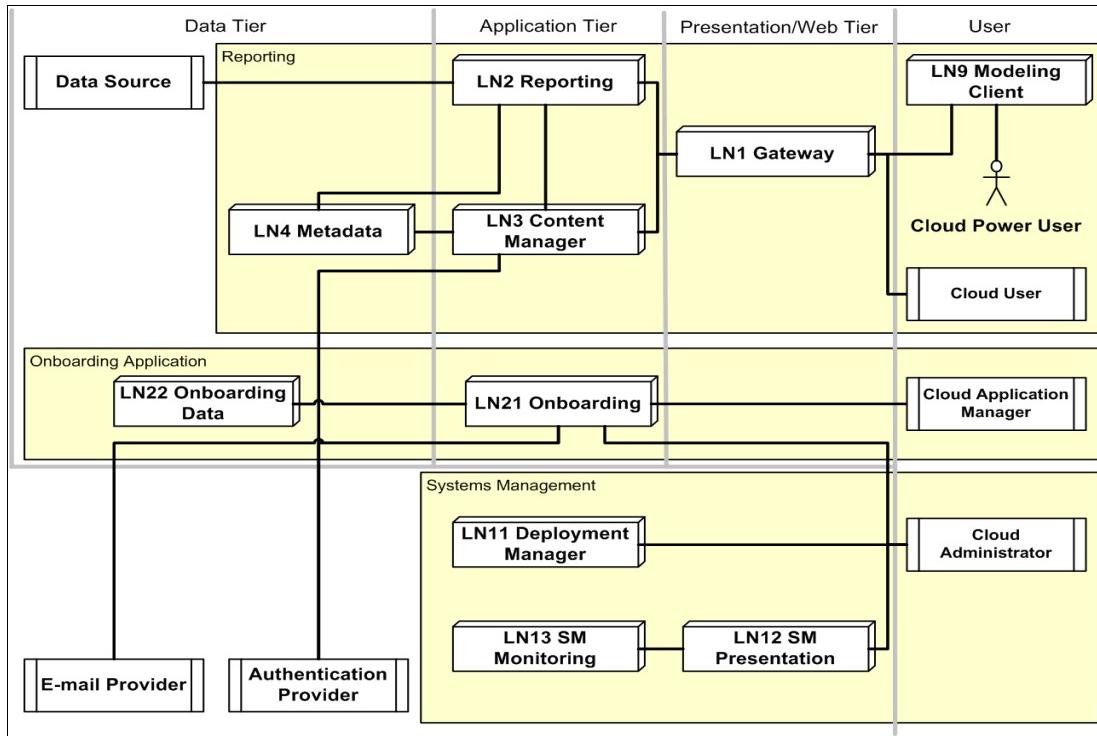
Figure 7-1   Logical operational model

7.1.2 Node description

In this section, we describe the nodes of the logical operational model. For each node we give an overview of the installed software and functional components and their non-functional requirement characteristics.

As a reference, we list the needed software products but provide more about them in section 7.3, "Physical operational model" on page 84 and in Part 4, "Implementation" on page 135.

LN1 Gateway

The gateway is the entry point for every user of the Smart Analytics Cloud. It manages all web communication for the IBM Cognos 8 Platform. The workload on the IBM Cognos 8 Gateway server is comparatively lightweight; therefore, it requires minimal processing resources. Because we are running a high-available cloud, we deploy multiple redundant gateways with an external HTTP load-balancing router.

The gateway consists of the following software:

▶ Web server: IBM WebSphere HTTP Server
▶ Application server: IBM WebSphere Application Server Network Deployment
▶ IBM Cognos Gateway
▶ IBM Tivoli Monitoring for O/S
▶ ITCAM for WebSphere
▶ ITCAM for HTTP Server

The non-functional requirements of this node are determined by the Cloud User actor.

LN2 Reporting

This node hosts the reporting engine of IBM Cognos 8 and the needed additional software and consists of the following software:

▶ Database management client: IBM DB2

▶ Application server: IBM WebSphere Application Server Network Deployment

▶ Reporting engine: IBM Cognos Report Server running on the application server

▶ IBM Cognos Dispatcher

▶ IBM Tivoli Monitoring for O/S

▶ ITCAM for WebSphere

The non-functional requirements of this node are determined by the node *LN1 Gateway*.

LN3 Content Manager

The main service of this node is to manage the storage of customer application data, including security, configuration data, models, metrics, report specifications, and report output. The Content Manager is needed to publish models, retrieve or store report specifications, handle scheduling information, and manage the Cognos name space. The Content Manager maintains its information in the databases on node *LN4 Metadata*. It consists of the following software:

▶ Database management client: IBM DB2

▶ Application server: IBM WebSphere Application Server Network Deployment

▶ Content management system: IBM Cognos Content Manager running on the application server

▶ IBM Tivoli Monitoring for O/S

▶ ITCAM for WebSphere

The non-functional requirements of this node are determined by the node *LN1 Gateway*.

LN4 Metadata

This node stores the queries and content of the reporting engine and the content management and consists of the following software:

▶ Database management system: IBM DB2
▶ Database for the queries
▶ Database for additional content
▶ IBM Tivoli Monitoring for DB2
▶ IBM Tivoli Monitoring for O/S

The non-functional requirements of this node are determined by the nodes *LN2 Reporting* and *LN3 Content Management*.

LN9 Modeling Client

This node is necessary to use the IBM Cognos 8 Framework software. It consists of the IBM Cognos Framework Manager software.

The non-functional requirements of this node are determined by the *Cloud Power User* actor.

LN11 Deployment Manager

This is an administrative node that manages all application servers of the application tier. It consists of the following software:

▶ Application server: IBM WebSphere Application Server Network Deployment
▶ IBM Tivoli Monitoring for O/S

The non-functional requirements of this node are determined by the necessity to change the application server configuration.

LN12 SM Presentation

The Tivoli Systems Management functions use a centralized presentation node for user access. It consists of the Portal: Tivoli Enterprise Portal Server (TEPS) software.

LN13 SM Monitoring

This node hosts a number of systems management components for managing the Smart Analytics Cloud. This node collects monitoring data from the agents on the other nodes and stores it in the monitoring database.

This node consists of the following software:

► Database management system: IBM DB2 Universal Database™
► Monitoring Server: Tivoli Enterprise Monitoring Server (TEMS)
► Monitor Warehouse: Tivoli Monitoring Data Warehouse (TDW)
► Provisioning: Tivoli Provisioning Manager (TPM)
► Directory: Tivoli Directory Server (TDS)

LN21 Onboarding

This node holds the necessary execution parts of the onboarding application. The onboarding application is a self-service application that gains access to the cloud and registers a new application into the cloud. The data is stored in the node *LN22 Onboarding Data*. It consists of the following software:

► Application server: IBM WebSphere Application Server Network Deployment
► Web forms framework and the onboarding application
► ITCAM for WebSphere

The non-functional requirements of this node are determined by the *Cloud Application Manager* actor.

LN22 Onboarding Data

This node holds the application data and metadata of the onboarding application.

It consists of the following software:

► Database management system: IBM DB2
► The database for web forms framework
► The database of the onboarding application
► IBM Tivoli Monitoring for DB2
► IBM Tivoli Monitoring for O/S

The non-functional requirements of this node are determined by the node *LN21 Onboarding*.

7.2 Discussion of non-functional requirements

Starting with the non-functional requirements (NFR) from 4.1, "Scope of the Smart Analytics Cloud" on page 28, we discuss possible elaborations of the logical operational model into a physical operational model. These NFRs are a sample but typical service level agreement for accessing a cloud with tight availability requirements. Depending on your environment, you might have a different set of requirements and can come to a different solution.

In this section, we discuss the following aspects:

► Availability
► Scalability
► Performance
► Multiple deployment environments

7.2.1 Availability

We have various types of users for this system with differing availability requirements:

► The *Cloud Application Manager* works with his requests (see section "Application" on page 70), which includes submitting a new request, changing the request, and accessing the status of his request.

These functions can be considered not business critical because they are only needed in the onboarding process to join the cloud. If the application is not available, this process might halt but does not break.

In general, the Cloud Application Manager's availability requirements do not imply what you will need for the affected nodes (*LN21 Onboarding* and *LN22 Onboarding Data*) to standby with session persistence. Just make sure that you can bring the application back in a timely manner.

► The *Cloud Administrator* has two responsibilities. He is involved with the onboarding process and has at least the same requirements as the *Cloud Application Manager* when accessing the onboarding application. Additionally, he manages the hardware, software, and configuration of the cloud.

His availability requirements imply that you have at least some kind of continuous operation so that this actor can always manage the cloud.

Affected nodes are responsible for systems management functionality: *LN11 Deployment Manager*, *LN12 SM Presentation*, and *LN13 SM Monitoring*.

► The *Cloud User* works with the reports. His availability requirements are potentially high. It is the nature of a cloud to be available when the end user wants to use it. Depending on the geographic distribution of your users and their typical working time, the requirements might even imply a 24x7 high availability for the system.

Consider building a high-available solution by clustering the logical nodes when creating physical nodes. The affected nodes are *LN1 Gateway, LN2 Reporting, LN3 Content Manager*, and *LN4 Metadata*.

> **Note:** The report execution depends on the underlying source systems that deliver the data to the Smart Analytics Cloud. Therefore the availability of the whole system depends heavily on the availability of the used source systems, which might not be in your control. You can only guarantee that the cloud is ready to perform the analytics.

► The *Cloud Power User* works with his workstation, which hosts the node *LN9 Modeling Client*. He has comparable availability requirements as the other management actors, maybe less than the Cloud Administrator.

The node *LN9 Modeling Client* is needed to manage some metadata but operates outside of the responsibility of the Smart Analytics Cloud. The entry point for this node (and therefore for the actor) into the cloud solution is the node *LN1 Gateway*. This entry can reuse the same infrastructure as the *Cloud User* or has its own physical node.

7.2.2 Scalability

The solution must respond to two basic scalability requirements:

► Additional users:

Only for the biggest user group (*Cloud Users)* can you expect a big increase over time, up to all employees in your enterprise. The number of managing users (*Cloud Power User*, *Cloud Application Manager*, or *Cloud Administrator*) will be quite small and probably will not expand much.

All *Cloud Users* must go through *LN1 Gateway*. This node can be scaled vertically by adding more processing power or horizontally by adding more physical nodes in a cluster. IBM Cognos 8 allows both methods, and this can be easy achieved. Because this node has high-availability requirements, you can start right from the beginning with a cluster to fulfill both NFRs.

All *Cloud Users* need access to *LN3 Content Manager* and *LN4 Metadata* to access the metadata for their reports. Both nodes can be scaled vertically by increasing the processing power.

► Additional workload because of additional users, additional reports, and more demanding existing reports.

The main report processing consists of two parts: executing SQL commands by the external systems and collecting the data and compiling the reports within the reporting engine on LN2 Reporting.

The source systems are responsible for executing SQL commands. *LN2 Reporting* can be scaled vertically by increasing the processing power and horizontally by adding additional nodes to a cluster. IBM Cognos 8 allows

both methods. Because this node has high-availability requirements, you can start from the beginning with a cluster to fulfill both NFRs.

7.2.3 Performance

There are two areas to look into when considering performance:

► End user performance

GUI-performance is driven by the performance of the underlying web and application servers and the connectivity to the end-users.

► Report performance

This is a crucial aspect for the Smart Analytics Cloud and influenced by a number of factors. In the next section, we discuss report performance as is typical for any kind of enterprise integration application that is modified for the Smart Analytics Cloud.

Report performance

As mentioned in the discussion about the scalability requirements, the report processing consists of two parts: query execution and report compilation. To drive performance, you might need to adjust both parts, as you do when considering scalability.

> **Note:** You must carefully manage the user expectations because some of the factors that drive performance lie in the responsibility of the source systems providing the data and not the Smart Analytics Cloud working with the data.

Query execution is determined by a number of factors:

► General performance of the source system

Because the source systems might not be under your control, you might have little chance to change the system. When estimating the performance of the overall cloud, look on the source systems, for example, do the source systems scale to the additional workload created by the cloud?

► Performance of data transport from the source system to the cloud

A similar factor is the performance of the source systems applies to the connection between the cloud and the source systems. The connection performance is affected by the network speed and the amount of data that is transported to the reporting engine. Of course the more data you must transport, the more time it will take. If the data cannot be transported fast enough, the overall performance can look poor, although it happens because of poor report creation.

▶ SQL execution, optimization, and tuning of the database management systems and the referenced tables

A central factor in the whole report processing lies in the execution of the SQL statements. A tuned database, the right table statistics, and well-optimized SQL are fundamental for the SQL performance—as always when dealing with databases. This time the performance depends on the applications that joined the cloud and on the skills of the report developer.

Report compilation is determined by:

▶ The size and the complexity of the reports: How much data is compiled into one report? It depends on the applications that joined the cloud and the design of the report developer.

▶ Sizing of the cloud: As mentioned for the scalability, this must not be an issue. With Cognos you can build a high-scalable solution.

7.2.4 Multiple deployment environments

In section 5.4.2, "Operational overview for multiple deployment environments" on page 61, we already showed you a sample operational overview for a multi environment Smart Analytics Cloud implementation. In this section, we drill deeper into the requirements of each environment and discuss different solution alternatives.

It is likely that you will need different cloud environments for:

▶ An environment for developing the reports
▶ An environment for tests and production readiness reviews
▶ An environment for the production

You might need additional environments. These environments will have different sets of non-functional requirements and will have a different number of users; however, all of these environments will share the same onboarding application.

Development environment

A development environment needs only minimal configuration. You can set up an environment for each application that joins the cloud.

There are a number of aspects to consider when dealing with development environments:

▶ Sometimes the developers need (or want) more access-rights on the hardware and software.

- A development environment might only be created when it is needed and disposed shortly after going into production. When a new environment is needed, it is created on-the-fly.

- Of course you might also share development environments for more than one application, just as you in a production environment.

All logical nodes can be placed on two physical machines to separate the application and the data parts:

- Application server consisting of: *LN1 Gateway, LN2 Reporting*, and *LN3 Content Manager*.

- Data server consisting of: *LN4 Metadata* and some databases for the development (as a replacement for the system actor Data Source).

The systems management functionality is probably centralized for all development systems.

Test environment

A test environment is always an in-between of development and production. It requires the same or comparable processing power as a productive environment but might not have the same availability requirements.

You will probably have one test environment for all applications to test the go-live-process and to test how the applications integrate in a common environment.

You can cluster the reporting *LN2 Reporting* and the gateway *LN1 Gateway* to have the necessary power, but do not consider high availability and skip standby systems.

Production environment

The production environment is a full-blown environment where you run all applications in one environment to take full advantage of the scalability features of z/VM and Linux on System z.

7.3 Physical operational model

In our lab environment, we simulate a high-available and highly-scalable production system. Therefore, to get a proper set of physical nodes, we duplicated the major logical nodes for use in a cluster or as a hot standby.

We made some additional important changes:

▶ Because we are currently not dealing with a lot of data, we put all logical nodes that are operating with databases on one physical node, which will have a hot standby.

▶ In our lab environment, we use an external provisioning provider; therefore, we must introduce a new system actor, *Provisioning Provider*, and will remove the provisioning from our physical nodes and include an interface to the actor.

▶ We do not integrate our onboarding application with an existing email infrastructure. We skip the *system actor Email Provider* and deactivate the functionality within our onboarding application.

In Table 7-2, we map to physical nodes.

Table 7-2   Mapping logical nodes onto physical nodes

Logical node	Physical node
LN1 Gateway	Multiple nodes in a cluster: ▶ PN1a Gateway 01 ▶ PN1b Gateway 02
LN2 Reporting	Multiple nodes in a cluster: ▶ PN2a Reporting 01 ▶ PN2b Reporting 02
LN3 Content Manager	Hot standby solution: ▶ PN3 Content Manager ▶ PN3s Content Manager Standby
LN4 Metadata LN22 Onboarding Data	Hot standby solution: ▶ PN4 Database Server ▶ PN3s Database Server Standby
LN9 Modeling Client	PN9 Modeling Client
LN11 Deployment Manager	PN11 Deployment Manager
LN12 SM Presentation	PN12 SM Portal Server
LN13 SM Monitoring	PN13 SM Monitoring But without the provisioning because its done by a new system actor *Provisioning Provider*
LN21 Onboarding	Multiple nodes in a cluster: ▶ PN21a Onboarding 01 ▶ PN21b Onboarding 02

All physical nodes run as a separate Linux on a System z guest in one z/VM LPAR. Figure 7-2 shows an overview of the resulting physical model.
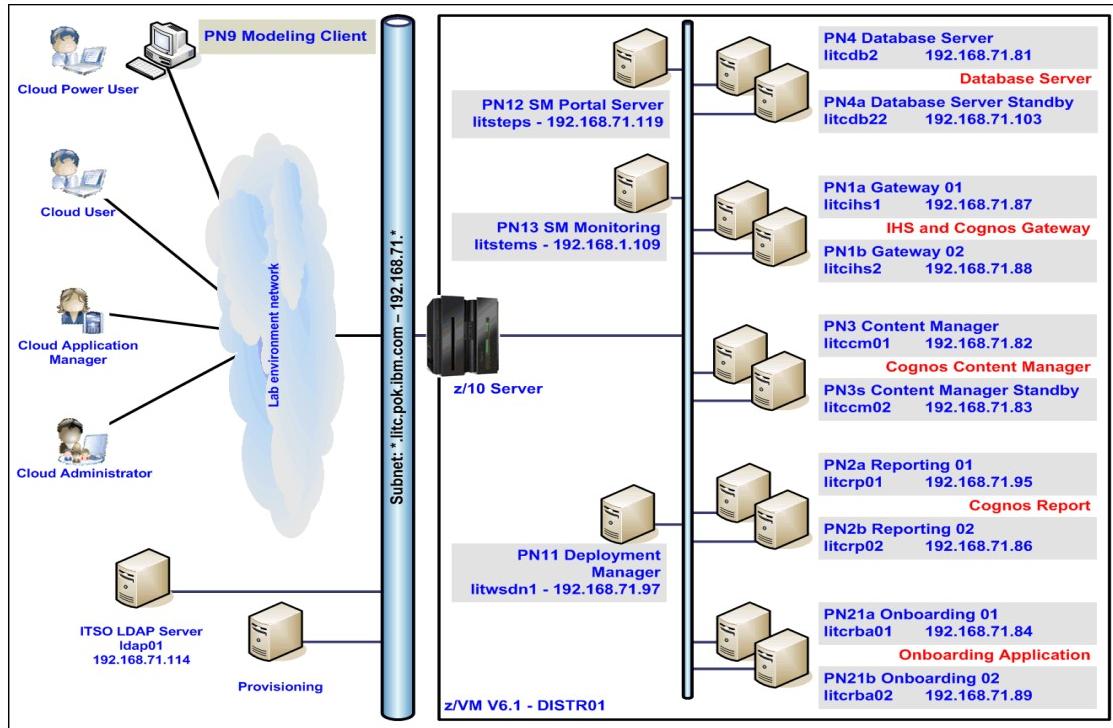

Figure 7-2   Physical Operational Model

7.3.1 Additional actors

An additional actor is the Provisioning Provider. In in our lab environment, we use an existing provisioning infrastructure, which is represented by this system actor.

Quantity: one

7.3.2 Node description

In this section, we describe each logical and physical node in the operational model shown in Figure 7-2.

PN1a Gateway 01 and PN1b Gateway 02

The *logical node LN1 Gateway* is placed on two physical nodes for scalability reasons. Table 7-3 on page 87 shows the PN1a Gateway 01 and PN1b Gateway 02 requirements.

Table 7-3   PN1a Gateway 01 and PN1b Gateway 02

Hardware	Linux guest in a z10 server ▶ 2 CPU ▶ 1 GB memory ▶ 15 GB disk space
Operating System	SUSE Enterprise Linux 11 (s390x 64bit)
Software	▶ IBM WebSphere HTTP Server V7.0.0.9 (64bit) ▶ IBM WebSphere Application Server Network Deployment V7.0.0.9 (64bit) ▶ IBM Cognos Gateway V8.4.1 (64bit)
Presentation	Cognos login
Data	None
Network connections	▶ http/https from PN9 browser of Cloud Application Manager ▶ SOAP/http to PN3, PN3s, PN2a, PN2b
Availability	As near to 24 x 7 x 52 as possible
Scalability	▶ The server can use more capacity by adding either extra capacity to existing engines on the server, by making the engines faster, or both. ▶ New server image can be added to the cluster to provide more Gateway server images
Security	▶ Users' authentication checking is done through user registry using LDAP ▶ Authorization checking is done through Cognos Access Manager
Systems management	▶ IBM Tivoli Monitoring for O/S ▶ ITCAM for WebSphere ▶ ITCAM for IBM HTTP Server
DR requirements	Yes

PN2a Reporting 01 and PN2b Reporting 02

The *logical node LN2 Reporting* is the centerpiece of the Smart Analytics Cloud; therefore, to show the scalability requirements, this node is placed on two

identical physical nodes. This set of physical nodes does the reporting within the cloud environment.

Table 7-4 on page 88 contains PN2a Reporting 01 and PN2b Reporting 02.

Table 7-4   PN2a Reporting 01 and PN2b Reporting 02

Hardware	Linux guest in a z10 server ► 2 CPU ► 4 GB memory ► 24 GB disk space
Operating System	SUSE Enterprise Linux 11 (s390x 64bit)
Software	► IBM DB2 Client V9.5 ► IBM WebSphere Application Server Network Deployment V7.0.0.9 (64bit) ► IBM Cognos Dispatcher (64bit) ► IBM Cognos Reporting V8.4.1 (64bit)
Presentation	Cognos report view
Data	None
Network connections	SOAP/http from PN1a, PN1b DB2 to PN4
Availability	As near to 24 x 7 x 52 as possible
Scalability	► The server can utilize more capacity by adding either extra capacity to existing engines on the server, by making the engines faster, or both ► New server image can be added to the cluster to provide more Report server images
Security	► Users' authentication checking is done through the user registry using LDAP ► Authorization checking is done through Cognos Access Manager
Systems management	► IBM Tivoli Monitoring for O/S ► ITCAM for WebSphere
DR requirements	Yes

PN3 Content Manager and PN3s Content Manager Standby

The logical node LN3 Content Manager is placed on two physical nodes: one as the primary node and the other as a standby.

Table 7-5 on page 89 contains PN3 Content Manager and PN3s Content Manager Standby.

Table 7-5   PN3 Content Manager and PN3s Content Manager Standby

Hardware	Linux guest in a z10 server ▶ 2 CPU ▶ 4 GB memory ▶ 24 GB disk space
Operating System	SUSE Enterprise Linux 11 (s390x 64bit)
Software	▶ IBM DB2 Client V9.5 ▶ IBM WebSphere Application Server Netwrok Deployment V7.0.0.9 (64bit) ▶ IBM Cognos Dispatcher (64bit) ▶ IBM Cognos Content Manager V8.4.1 (64bit)
Presentation	None
Data	None
Network connections	▶ SOAP/http from PN1a, PN1b ▶ JDBC to PN4
Availability	As near to 24 x 7 x 52 as possible
Scalability	▶ The server can utilize more capacity by adding either extra capacity to existing engines on the server, by making the engines faster, or both. ▶ New server image can be added to the cluster, but it can only be running as Standby mode.
Security	▶ Users' authentication checking is done through the user registry using LDAP. ▶ Authorization checking is done through Cognos Access Manager.
Systems management	▶ IBM Tivoli Monitoring for O/S ▶ ITCAM for WebSphere
DR requirements	Yes

PN4 Database Server and PN4s Database Server Standby

In our reference installation, we do not have a number of external source systems; instead, we built a sample database based on all of the external source systems.

The physical node combines the following elements:

▶ LN22 Onboarding Data
▶ LN4 Metadata
▶ System Actor *Data Source*

Table 7-6 contains the PN4 Database Server and PN4s Database Server Standby.

Table 7-6   PN4 Database Server and PN4s Database Server Standby

Hardware	Linux guest in a z10 server ► 2 CPU ► 4 GB memory ► 30 GB disk space
Operating System	SUSE Enterprise Linux 11 (s390x 64bit)
Software	IBM DB2 V9.5 (64bit)
Presentation	None
Data	► Onboarding application data ► Cognos Server Content Store ► Sample source data
Network Connections	► JDBC to PN3, PN3s, PN21a, PN22b ► DB2 to PN2a, PN2b
Availability	As near to 24 x 7 x 52 as possible
Scalability	The server can utilize more capacity by adding either extra capacity to existing engines on the server, by making the engines faster, or both.
Security	Users' authentication checking is done through the native O/S.
Systems management	► IBM Tivoli Monitoring for O/S ► IBM Tivoli Monitoring for DB2
DR requirements	Yes

PN9 Modeling Client

This node hosts the logical node LN9 Modeling Client. Table 7-7 contains the requirements for the PN9 Modeling Client.

Table 7-7   PN9 Modeling Client

Hardware	WINDOWS compatible PC
Operating System	WINDOWS ► 1GB memory (min)

Software	► IBM Cognos Framework Manager V8.4.1 ► Microsoft® Data Access Component (MDAC) V2.6 or higher ► DB2 Connect™ V9.5
Presentation	To actor *Cloud Power User*
Data	None
Network connections	http/https to PN1a, PN1b
Availability	N/A
Scalability	N/A
Security	Users' authentication checking is done through native OS
Systems management	N/A
DR requirements	N/A

PN11 Deployment Manager

The *logical node L11 Deployment Manager* becomes one physical node.

It can put it on one of the other nodes that host a WebSphere Application Server, but every instance must be treated equally. Table 7-8 contains the requirements for the PN11 Deployment Manager.

Table 7-8   PN11 Deployment Manager

Hardware	Linux guest in a z10 server ► 2 CPU ► 2 GB memory ► 16 GB disk space
Operating System	SUSE Enterprise Linux 11 (s390x 64bit)
Software	IBM WebSphere Application Server Network Deployment V7.0.0.9 (64 bit)
Presentation	WebSphere Application Server Admin Console
Data	None
Network connections	SOAP/RMI to PN1a, PN1b, PN2a, PN2b, PN3, PN3s, PN21a, PN22b
Availability	As near to 24 x 7 x 52 as possible

Scalability	The server can utilize more capacity by adding either extra capacity to existing engines on the server, by making the engines faster, or both.
Security	Users' authentication checking is done through user registry using LDAP
Systems management	IBM Tivoli Monitoring for O/S
DR requirements	Yes

PN12 SM Portal Server

The *logical node LN12 SM Presentation* becomes one physical node. Table 7-9 contains the requirements for the PN12 SM Portal Server.

Table 7-9   PN12 SM Portal Server

Hardware	Linux guest in a z10 server ► 2 CPU ► 1 GB memory ► 16 GB disk space
Operating System	SUSE Enterprise Linux 11 (s390x 64bit)
Software	► IBM DB2 V9.5 (64bit) ► Tivoli Enterprise Portal Server
Presentation	http/https from browser of Cloud Administrator
Data	None
Network connections	IP.PIPE to PN10
Availability	As near to 24 x 7 x 52 as possible
Scalability	The server can utilize more capacity by adding either extra capacity to existing engines on the server, by making the engines faster, or both
Security	Users' authentication checking is done through user registry using LDAP
Systems Management	N/A
DR requirements	Yes

PN13 SM Monitoring

The *logical node LN13 Systems Management Monitoring* becomes on physical node. Table 7-10 contains the requirements for PN13 SM Monitoring.

Table 7-10   PN13 SM Monitoring

Hardware	Linux guest in a z10 server ▶ 2 CPU ▶ 1 GB memory ▶ 9 GB disk space
Operating System	SUSE Enterprise Linux 11 (s390x 64bit)
Software	▶ IBM DB2 V9.5 (64bit) ▶ Tivoli Enterprise Monitoring Server ▶ Tivoli Monitoring Datawarehouse ▶ Tivoli Directory Server
Presentation	Interface to the Provisioning Provider
Data	None
Network connections	IP.PIPE from PN1a, PN1b, PN2a, PN2b, PN3, PN3s, PN4, PN4s, PN11, PN21a, PN21b
Availability	As near to 24 x 7 x 52 as possible
Scalability	The server can utilize more capacity by adding either extra capacity to existing engines on the server, by making the engines faster, or both
Security	Users' authentication checking is done through user registry using LDAP
Systems management	N/A
DR requirements	Yes

PN21a Onboarding 01 and PN21b Onboarding 02

The *logical node LN21 Onboarding* is placed on two physical nodes in a cluster to show how to cope with high-availability requirements for this node. Normally, that is not necessary because the non-functional requirements are not that high for this kind of operation. Table 7-11 on page 94 contains the PN21a Onboarding and PN21b Onboarding Standby.

Table 7-11   PN21a Onboarding and PN21b Onboarding Standby

Hardware	Linux guest in a z10 server ▶ 2 CPU ▶ 2 GB memory ▶ 30 GB disk space
Operating System	SUSE Enterprise Linux 11 (s390x 64bit)
Software	▶ IBM WebSphere Application Server Network Deployment V7.0.0.9 (64bit) ▶ Web forms framework ▶ Onboarding Application
Presentation	
Data	None
Network connections	http/https to Cloud Application Manager JDBC to PN4 Database Server
Availability	As near to 24 x 7 x 52 as possible
Scalability	▶ The server can utilize more capacity by adding either extra capacity to existing engines on the server, by making the engines faster, or both. ▶ New server images can be added to the cluster to provide more application server images
Security	Users' authentication checking is done through user registry using LDAP
Systems management	▶ IBM Tivoli Monitoring for O/S ▶ ITCAM for WebSphere
DR requirements	Yes

7.4 Walkthrough

The nodes in the operational model communicate through a variety of paths. Figure 7-3 on page 95 shows the key communication paths. Figure 7-3 on page 95 does not represent the complete end-to-end flow of a request because the flow varies for each of the many possible request types.

For more information about specific request flows, see the IBM Cognos 8 Business Intelligence Architecture and Deployment Guide at:

http://download.boulder.ibm.com/ibmdl/pub/software/data/cognos/document ation/docs/en/8.4.0/crn_arch.pdf
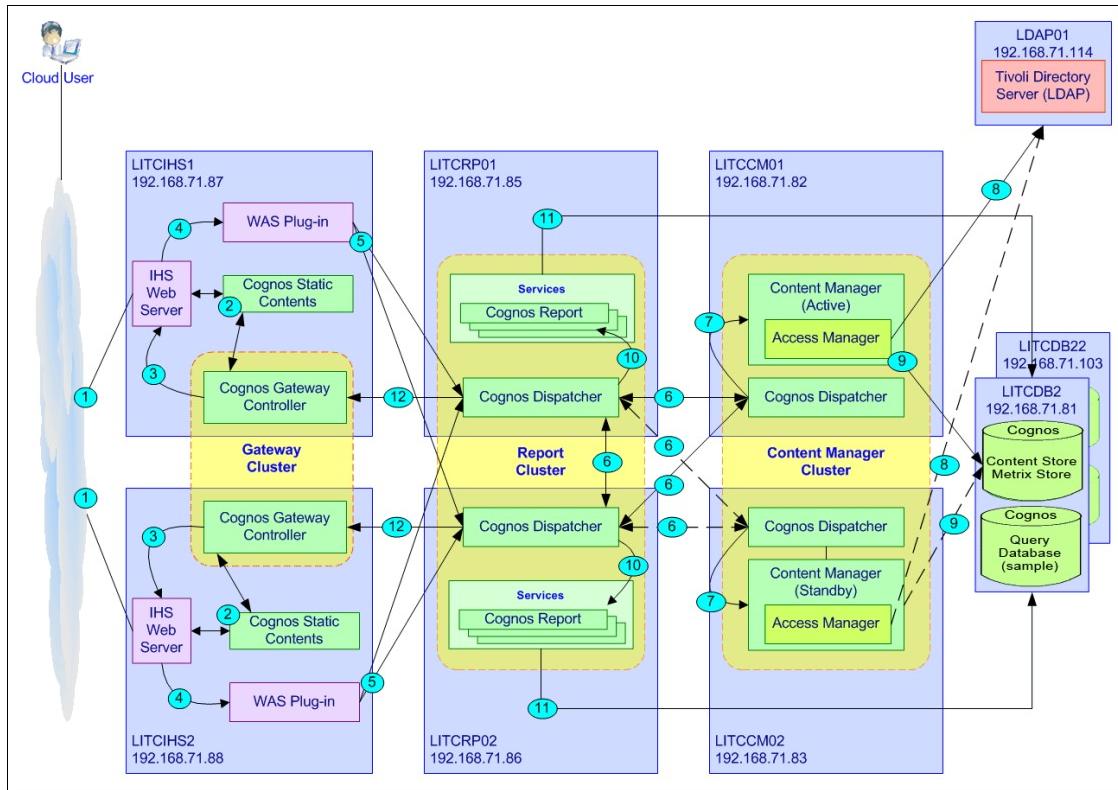
Figure 7-3   Communication flow in the Cognos environment

The key communication points are:

1. User browser sessions connect to the cloud through one of two Web servers.

2. The Web server sends requests to the Cognos gateway and returns information from the gateway to the browser.

3. The gateway is configured to use URI of the WebSphere Application Server plug-in for dispatching. When the gateway needs to communicate with a dispatcher, the request goes to the Web server.

4. The Web server directs gateway dispatcher requests to the WebSphere Application Server plug-in.

5. The WebSphere Application Server plug-in is aware of all dispatchers in the report cluster and automatically load balances requests to dispatchers.

6. The dispatcher can pass requests to the Content Manager or to other dispatchers in the Application Tier, as needed. The dispatchers are aware of all content managers and can direct requests to the standby content manager, if the primary is unavailable.

7. The dispatcher on the Content Manager can send requests to the appropriate services on the server.

8. The Access Manager authenticates requests with the LDAP server.

9. The Content Manager stores or retrieves information from the content store database.

10. The dispatcher communicates with the report services and other services in the application tier.

11. The report service connects to the query database to run reports.

12. The dispatcher sends information to the browser using the gateway.

8

Cloud Management architecture

In this chapter, we discuss architecture considerations for the system management of the Smart Analytics Cloud. We describe the IBM approach for cloud system management embedded in an holistic cloud reference architecture along with the structure of the support services. We also highlight some of the important support services in a Smart Analytics Cloud. Finally, we show the technical solutions architecture for some of the support services.

97

8.1 Cloud management architecture

In Chapter 5, "Architecture overview" on page 47, Chapter 6, "Functional architecture" on page 63, and Chapter 7, "Operational architecture" on page 73, we described the functional and operational aspects of the application architecture that is required for the analytics components of the Smart Analytics Cloud.

In this chapter, we round the architecture up, by introducing, as a starting point, a reference architecture for cloud computing systems management and take a dive into some important aspects.

8.1.1 Architecture overview

As discussed in earlier parts of this book, such as section 3.2, "Cloud computing" on page 21, there is more to a cloud than just the software and hardware components. This can also be seen in Figure 8-1 on page 100, which shows a typical cloud management architecture overview. It consists of three major building blocks based on the high-level roles that are involved in cloud computing:

► Cloud Service Consumer
► Cloud Service Provider
► Cloud Service Developer

Cloud Service Consumer

Cloud Service Consumer includes every user and system that uses or consumes resources of the cloud. This includes the actors listed in section 5.3.3, "Data Sources" on page 58 and some of the roles described in Chapter 15, "Service life cycle" on page 249 and in section 16.1, "Roles in the onboarding process" on page 254.

Additional block we have:

► Cloud Service Integration Tools: For the Cloud Service Consumer, it is important to be able to integrate cloud services with their on-premise IT. This functionality is most relevant in the context of hybrid clouds, where the seamlessly integrated management and use of different cloud services and on-premise IT is critical.

► Consumer In-house IT: Besides IT capabilities that are consumed as cloud services, consumers of such IT will continue to have in-house IT, which can be managed in a traditional non-cloud fashion. In case the functionality of the existing in-house IT is integrated with cloud services that are consumed from

a cloud service provider, the aforementioned capability to integrate with external clouds is required.

Cloud Service Provider

The Cloud Service Provider offers cloud services that correspond to our "Smart Analytics Cloud functions" on page 56, which is implemented using Cognos 8 BI in our lab environment. In addition, the Cloud Service Provider supplies Infrastructure and a Common Cloud Management Platform. The Common Cloud Management Platform consists of two major building blocks:

► The Business Support Services (BSS) consists of the business-relevant platform services.

► The Operational Support Services (OSS) represent the more infrastructure and operational aspects of the Common Cloud Management Platform.

Cloud Service Developer

A Cloud Service Developer uses Service Development Tools to develop new cloud services, which includes both the development of runtime artifacts (for example, database persistence, transactional handling, and so on) and management-related aspects (for example, monitoring, metering, provisioning, and so on). In this context, the service development tools support the cloud service developer in creating a service template and a service offering, whereas the service template defines how the Common Cloud Management Platform (CCMP) OSS functionality is used in the context of the respective cloud service and the service offering specifies how the CCMP BSS functionality is used in the context of the respective cloud service.

In the context of a particular infrastructure or platform as-a-service offering, there might also be tooling to develop artifacts that are specific to the particular cloud service. In our lab environment, for example, it is necessary to use image creation tools for developing images that can be deployed in the cloud service.

So in summary, there are two categories of service development tooling: tooling to develop a cloud service by itself and tooling to develop artifacts that are specific to the respective cloud service.

Figure 8-1 on page 100 shows a Cloud management architecture overview.
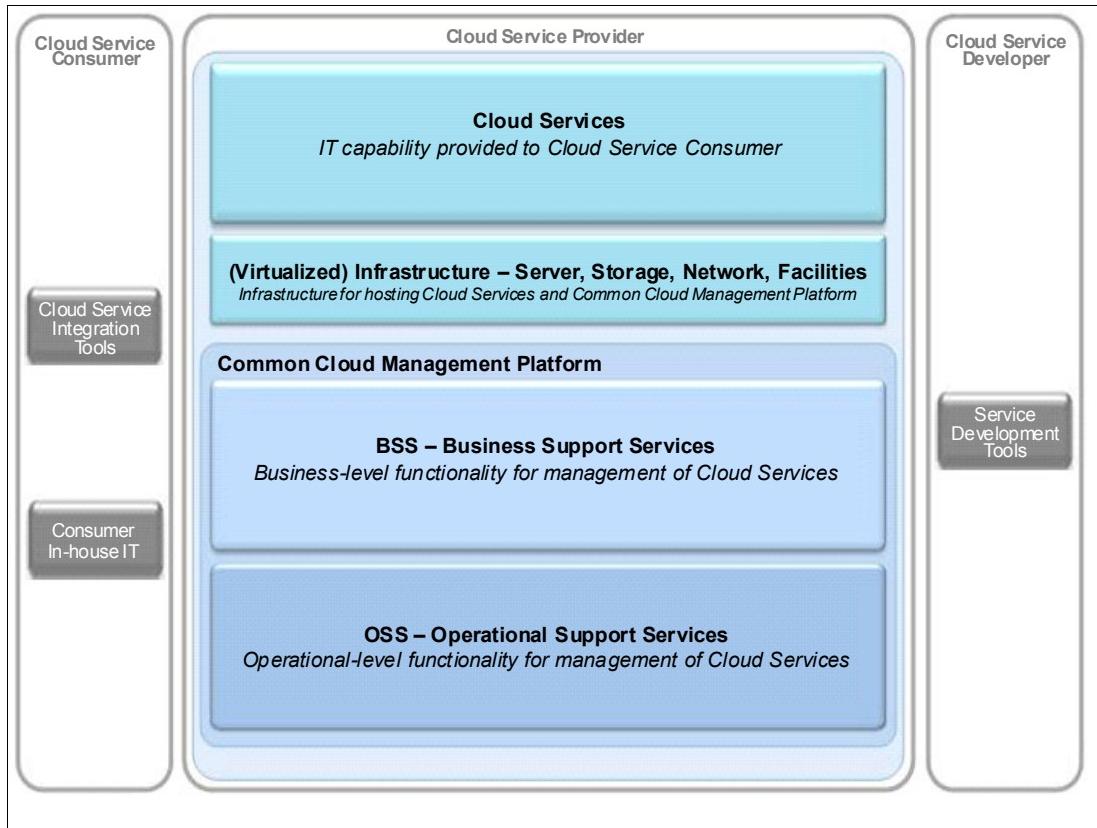
Figure 8-1   Cloud management architecture overview

In the next sections of this chapter, we take a deeper look into the Cloud Service Provider building block.

8.1.2 Cloud services

Cloud services represent any type of (IT) capability that is provided by the Cloud Service Provider to Cloud Service Consumers. Typical categories of cloud services are infrastructure, platform, software, or Business Process Services. In contrast to traditional (IT) services, cloud services have attributes associated with cloud computing, such as a pay-per-use model, self-service usage, flexible scaling, and shared underlying IT resources.

Figure 8-2 on page 101 shows the typical cloud services structure.

In our lab environment, the cloud service we deliver is a Cognos 8 BI environment.
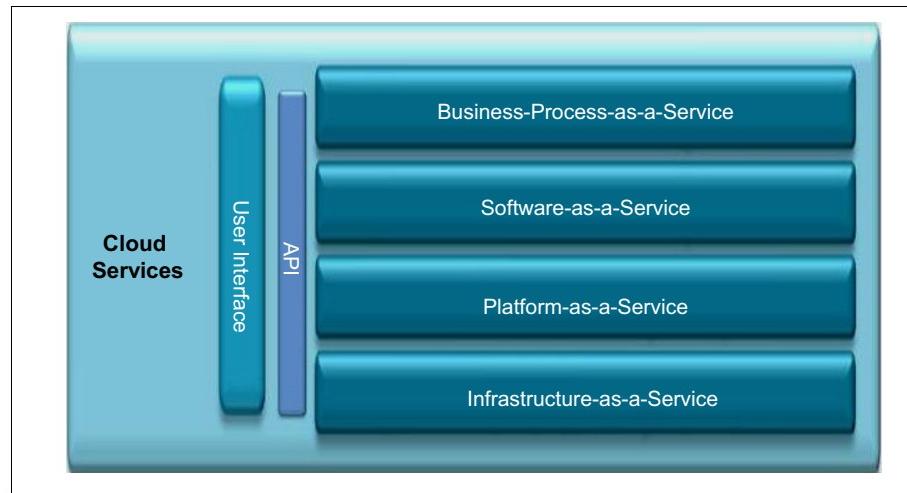

Figure 8-2   Cloud services

Cloud services can be built on top of each other, for example, a software service can consume a platform or infrastructure service as its basis, and a platform service can consume an infrastructure service as its foundation. In general, architectural principles postulate to share as much as possible across cloud services with respect to management platform and underlying infrastructure. However, it does not require one single, fully homogeneous infrastructure, which of course is the ideal goal, but given different infrastructure requirements, it is not possible, for example, if a particular cloud services has specific infrastructure needs, it is clearly allowed to run this cloud service on a dedicated infrastructure. In any case, each Cloud Service offered by a Cloud Service Provider is known to the BSS and OSS.

A cloud service provider offers cloud services as a result of conscious business decisions. Offering a cloud service to internal or external customers must be supported by a corresponding solid business model and investments for the development and operations of the cloud service.

Defining and delivering a cloud service requires nailing down all corresponding functional and non-functional requirements. The artifacts developed in addition to the cloud service have only minimal room to change how these functional and non-functional requirements are addressed. This is not to be viewed as something negative because it is a core value proposition of cloud services to provide strict guidelines with respect to how they can be exploited because this is the main factor driving a reduction in cost of artifact development. The easier it is

to develop artifacts for such a cloud service, the more likely the cloud service is successful.

8.1.3 Virtualized infrastructure

The virtualized infrastructure includes all infrastructure elements needed on the cloud service provider side to provide cloud services, which includes facilities, server, storage, and network resources and how these resources are wired up, placed within a data center, and so on. In case of virtualization, this also includes virtualization information, such as hypervisors. It does not include any virtualization management software because that is part of the virtualization management component of the "Operational Support Services" on page 105.

Figure 8-3 shows the components of a typical virtualized infrastructure.
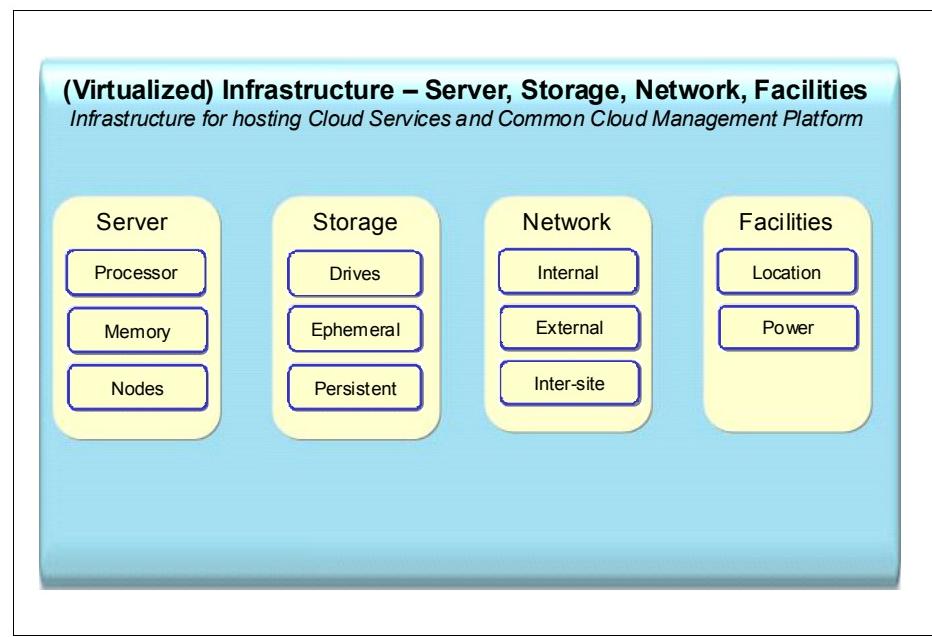
Figure 8-3   Virtualized infrastructure

The decision whether the infrastructure is virtualized or not depends on the actual workload characteristics to be run on the respective infrastructures. This is not a violation of the architectural principles postulating as much as possible commonality across cloud services. While maximum commonality is a core architectural principle, it is allowed to have different infrastructure architecture per workload category. However, a requirement in any case is that all of these

infrastructures get managed from a single, central CCMP and that CCMP can place instances of each cloud services on the corresponding infrastructure.

The more homogeneous the infrastructure is, the more it caters to the standardization needs of a cloud environment. Homogeneity on the infrastructure side is critical for enabling the high degrees of automation and economies of scale, which are the key characteristics of any cloud environment. However, it must be acknowledged that in many cloud deployments (specifically private clouds) there are different workloads to be provided as a cloud service, and each of these workloads might have special infrastructure needs. So although the ideal case is total homogeneity on the infrastructure side, it is important to note that there will be cloud installations that have some variants in the infrastructure elements (for example, different HW platforms).

The infrastructure is managed by the OSS as part of the CCMP, whereas the CCMP by itself also runs on the infrastructure.

> **Note:** The physical existence of a virtualized infrastructure on the cloud service provider side is not mandatory because a cloud service provider can consume infrastructure as a service (and the required CCMP) from a different cloud service provider and put higher-value cloud services on top.

8.1.4 Common Cloud Management Platform

The Common Cloud Management Platform (CCMP) contains a set of business and operational management focused services that must be used by Cloud Services to actually be a cloud service.

The CCMP is responsible for:

► Delivering instances of Cloud Services of any category to Cloud Service Consumers

► The ongoing management of all Cloud service instances from a provider perspective

► Allowing Cloud Service Consumers to manage their Cloud Service instances in a self-service fashion

The technical aspects of a Cloud Service are captured in a service template, which is also the artifact that describes how the OSS capabilities of the CCMP are exploited within the context of the respective Cloud Service. The considerations that we made for the provided services are described in 4.1, "Scope of the Smart Analytics Cloud" on page 28.

The CCMP is defined as a general purpose cloud management platform built to support the management of any category of cloud service.

As the name already implies, the CCMP is structured as a platform. Based on the platform nature, the CCMP contains a set of services (some of them optional) that are used within the context of a specific cloud service. The CCMP services provided to the cloud service developers must be distinguished from the cloud services developed by cloud service developers. Cloud Service developers are strongly encouraged to use the services provided by the CCMP to enable the economies of scale needed for achieving the extremely high degrees of efficiency associated with any cloud computing environment.

As an example, it is required to apply a special audit for any software component that has financial impact, which means the component that is doing billing for the consumption of a cloud service must be audited. By establishing a single deployment of a billing component, shared amongst multiple cloud services, the complex and time-consuming audit process must only be executed one time and can then be used for any number of cloud services instead of executing a separate audit each time a new cloud service is deployed in an environment without a CCMP. Clearly, this concept of sharing enables economies of scale and does not only apply to the billing service of BSS but also for any other management service that is part of a CCMP deployment.

The CCMP is split into two main elements:

▶ Operational Support Services
▶ Business Support Services

Business Support Services

BSS represents the set of business-related services exposed by the CCMP (billing, entitlement, invoicing, and so on), which must be exploited by Cloud Service Developers to take advantage of the common cloud management platform.

Figure 8-4 on page 105 shows details about the business support services for a cloud.

The BSS provides services that either enable the cloud service provider or facilitates certain task to deliver the cloud from a business perspective. It contains the services offering management, customer management, pricing and rating, order management, entitlement management, subscriber management, general accounting, invoicing, billing, peering and settlement, contract and agreement management, opportunity to order, metering, analytics and reporting, and service offer catalog. Besides the business aspect, there is also the technical side of cloud computing because it includes a fast provisioning of

standardized IT products and services. In the reference architecture, this area is called Operational Support Services.
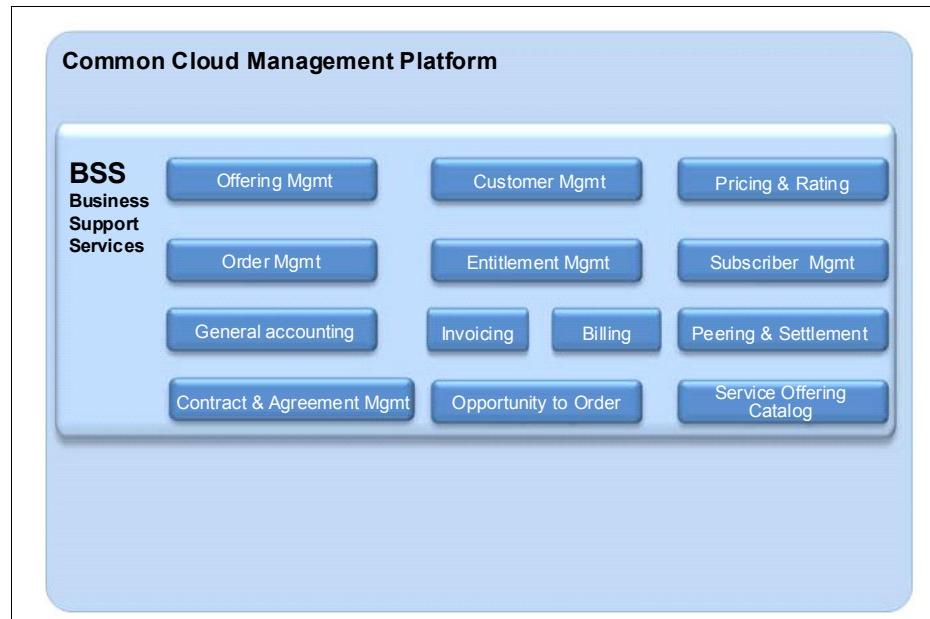

Figure 8-4   Cloud management architecture Business support services

Like any other component of the CCMP, the BSS is generic across all cloud service types and can be configured to behave appropriately in the context of the managed cloud services. As an example, the billing service of the CCMP BSS must be usable to do billing for the consumption of virtual machines (IaaS), a multi-tenancy capable middleware platform, such as the Common Cloud Service Platform (PaaS), and for collaboration services, such as LotusLive™ (SaaS). This drives the need for a proper platform-level definition of all BSS components and exploitation artifacts enabling cloud service developers to prime the behavior of each BSS component in a cloud-service specific fashion.

Operational Support Services

OSS represents the set of operational management and technical-related services exposed by the CCMP, which must be exploited by Cloud Service Developers to take advantage of the common cloud management platform.

Figure 8-5 on page 106 shows the components that are included in the operational support services for a cloud.

The OSS contains the following services: service delivery catalog, service template, service automation management, service request management,

change and configuration management, image life cycle management, provisioning, incident and problem management, IT service level management, monitoring and event management, IT asset and license management, capacity and performance management, and virtualization management.

Later in this chapter we discuss how we implemented the key services that are relevant for our lab environment. However, some of the services are key services but were not implemented in our environment. We provide a more detailed description of these and give an outlook on how they can be implemented in later releases of our lab environment.

Obviously, the ideal case from a cost optimization and economies-of-scale perspective is to use as much as possible shared CCMP OSS/BSS functionality, but if necessary other options are also viable. In general, OSS and BSS are viewed as the (integrated) set of management platform functionality underpinning the operation of a cloud service (similar to middleware being the functional / runtime platform). Figure 8-5 shows the operational support services in the cloud management architecture.
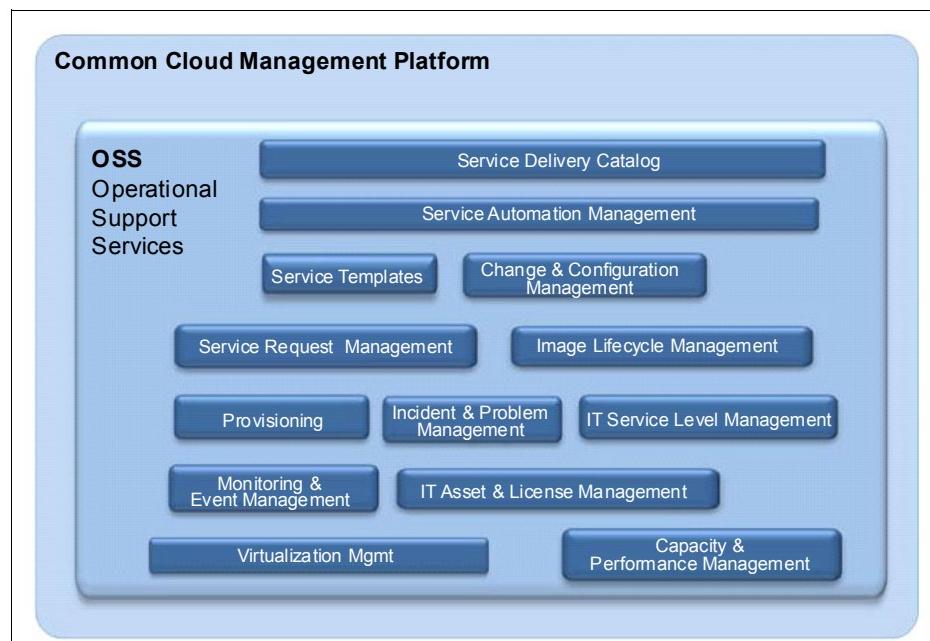

Figure 8-5   Cloud management architecture Operational support services

Many management domains shown in the OSS can also be encountered in traditionally managed data centers (monitoring and event management, provisioning, incident and problem management, and so on), while other components are new and rather specific to the degrees of automation and

efficiency that are associated with clouds (service automation, image life cycle management).

Particularly for the traditional management domains it is important to note that conceptually they are the same in the cloud world and in the traditional world, whereas in a cloud world these domains are implemented in radically different ways taking advantage of the high degrees of homogeneity in a cloud, for example, a traditionally managed data center is implemented in a way that an incident gets raised if a physical server fails, a ticket gets opened, and assigned to an administrator (maybe 2 AM in the morning). After some time, an escalation takes place if the administrator has not resolved the ticket until then. In contrast, in a cloud environment, there is also incident and problem management, whereas here a broken physical machine can be left broken on the floor until some later point of time because the virtual machines that are running on that physical machine can be brought up on another one. Both scenarios address incident and problem management, but in radically different ways and for radically different labor costs. A similar *cloudyfied* perspective exists also for most other OSS components.

The platform notion of CCMP obviously also applies to all components defined as part of the OSS: A proper platform-level definition of all OSS components and exploitation artifacts enabling cloud service developers is needed to prime the behavior of each BSS component in a cloud service specific fashion.

8.1.5 Security and resiliency

Security and resiliency are cross-cutting aspects spanning the CCMP, the (virtualized) infrastructure, and Cloud Services. Both of these non-functional aspects must be viewed from an end-to-end perspective, including the security set up of CCMP by itself, the way the infrastructure is set up (in terms of isolation, network zoning set up, and data center set up for disaster recovery) and how the cloud services are structured. More about in section 8.6, "Security enforcement and management" on page 123.

8.2 OSS: Monitoring and event management

Having introduced the architecture overview for cloud computing, we now discuss the implementation approach we chose in our lab environment for the required OSS and BSS components.

In this section, we discuss the functions and the solutions of the Operational Support Services for monitoring and event management.

8.2.1 SAC monitoring and event functions

At its most basic level, monitoring is the process of defining what elements of the IT infrastructure are expected to fail and then putting mechanisms in place to notify a person or a process when the risks of failure are near. Monitoring is therefore an analytical process that is built up of management patterns over time. IT Systems are increasingly complex and therefore the art of monitoring is proactively checking those critical IT elements in the system without getting into micro management of systems that in itself causes considerable destabilizing load on the system monitored and other infrastructure management components.

In the context of the Smart Analytic Cloud, the important concerns for monitoring and event that need to be addressed are:

► Centrally manage all monitored servers and events.

► Manage operating systems, databases, applications, and servers in general in the wide variety of software and hardware deployed in the cloud and also:

– Hypervisor availability and performance monitoring

– Extensions to event management that correlates network, server hardware, local storage, network storage, network elements, operating system, and hypervisor events

► Warehouse all events along with any gathered correlation metadata.

► Report all warehouse events under one roof (performance, events, event actions).

► Provide a common, flexible, and easy-to-use browser interface and customizable displays to facilitate system monitoring.

► Detect and recover potential problems in essential system resources automatically.

► Include, as part of the system monitoring software package, easy-to-use warehouse and advanced reporting capability.

► Provide situation-based alerts and real-time and historic reports.

8.2.2 SAC monitoring solution with Tivoli Monitoring

In our Smart Analytics Cloud, we choose to use IBM Tivoli Monitoring software because it fulfils the functions described in the preceding section.

IBM Tivoli Monitoring proactively monitors essential system resources, detects bottlenecks and potential problems, and automatically responds to events. Proactive system monitoring often identifies problems early, enabling rapid fixes

before end users experience significant impact to their performance. It has immediate best practices for identifying and resolving infrastructure problems. The data collected with Tivoli Monitoring can also be used to drive timely performance and capacity planning activities to avoid outages from resource over-utilization.

Figure 8-6 shows the Tivoli Monitoring component model.
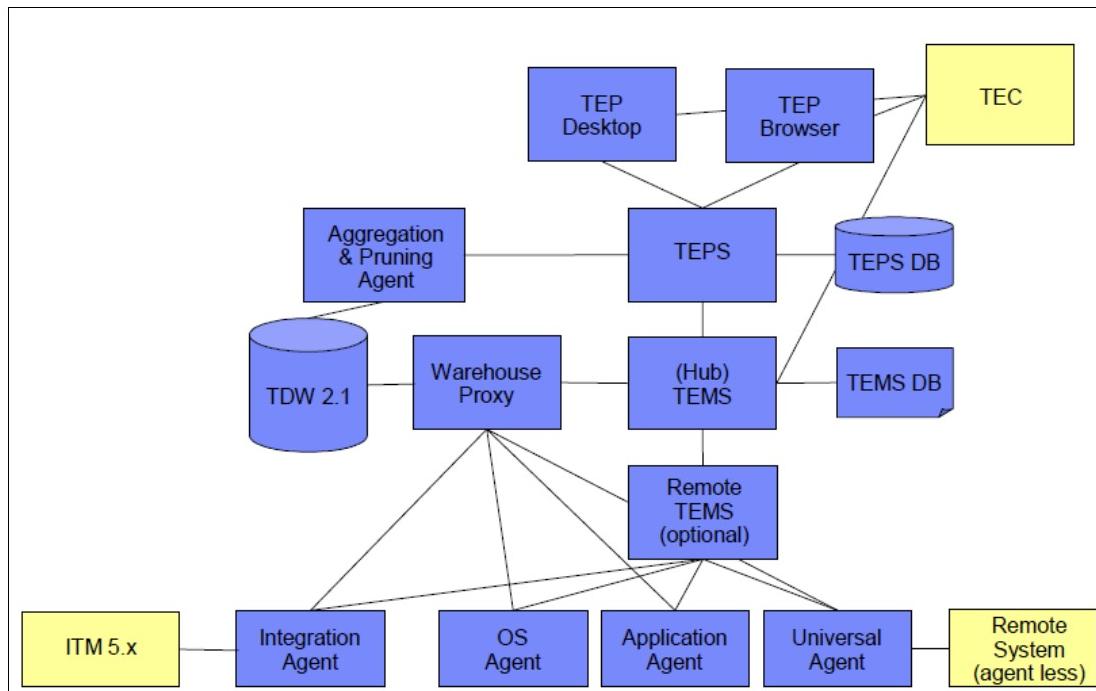
Figure 8-6   Tivoli Monitoring component model

The Tivoli Enterprise Monitoring Server (referred to as the monitoring server) is the initial component to install to begin building the IBM Tivoli Monitoring Services infrastructure. It is the key component on which all other architectural components depend directly. The TEMS acts as a collection and control point for alerts received from agents, and collects their performance and availability data.

The primary TEMS is configured as a Hub(*LOCAL). All IBM Tivoli Monitoring installations require at least one TEMS configured as a Hub. Additional Remote(*REMOTE) TEMS can be installed later to introduce a scalable hierarchy into the architecture.

The Tivoli Enterprise Portal Server (referred to as the portal server) is a repository for all graphical presentation of monitoring data. The portal server

database also consists of all user IDs and user access controls for the monitoring workspaces. The TEPS provides the core presentation layer, which allows for retrieval, manipulation, analysis, and preformatting of data. It manages this access through user workspace consoles.

The TEP client (referred to as the portal client) is a Java-based user interface that connects to the Tivoli Enterprise Portal Server to view all monitoring data collections. It is the user interaction component of the presentation layer. The TEP brings all of these views together in a single window, so you can see when any component is not working as expected. The client offers two modes of operation: a Java desktop client and an HTTP browser.

The agents (referred to as managed systems) are installed on the system or subsystem requiring data collection and monitoring. The agents are responsible for data gathering and distributing attributes to the monitoring servers, including initiating the heartbeat status. These agents test attribute values against a threshold and report these results to the monitoring servers. The TEP displays an alert icon when a threshold is exceeded or a value is matched. The tests are called situations.

The Warehouse Proxy is a unique agent that performs only one task, collecting and consolidating all Historical Data Collections from the individual agents to store in the Tivoli Data Warehouse.

The Summarization and Pruning agent is a unique agent that performs the aggregation and pruning functions for the historical raw data on the Tivoli Data Warehouse.

The Tivoli Data Warehouse is the database storage that contains all of the Historical Data Collection.

For further information see *Getting Started with IBM Tivoli Monitoring 6.1 on Distributed Environments*, SG24-7143.

ITM monitors all components of the SAC with agents deployed on each of them (they refer to the application agent in the component model):

► ITCAM for WebSphere agents monitor the Cognos servers running on top of WebSphere Application Server.

► ITCAM for HTTP agents monitor the Cognos components running on top of the IBM HTTP Server.

► ITM for Database agents monitor the Cognos databases.

- ► Operating System agents monitor all the Linux operating systems in the Cloud.
- ► OMEGAMON XE on z/VM agents monitor the underlying virtualization layer executed by z/VM.

Figure 8-7 shows the different agents deployed on the Linux guests of or lab environment.
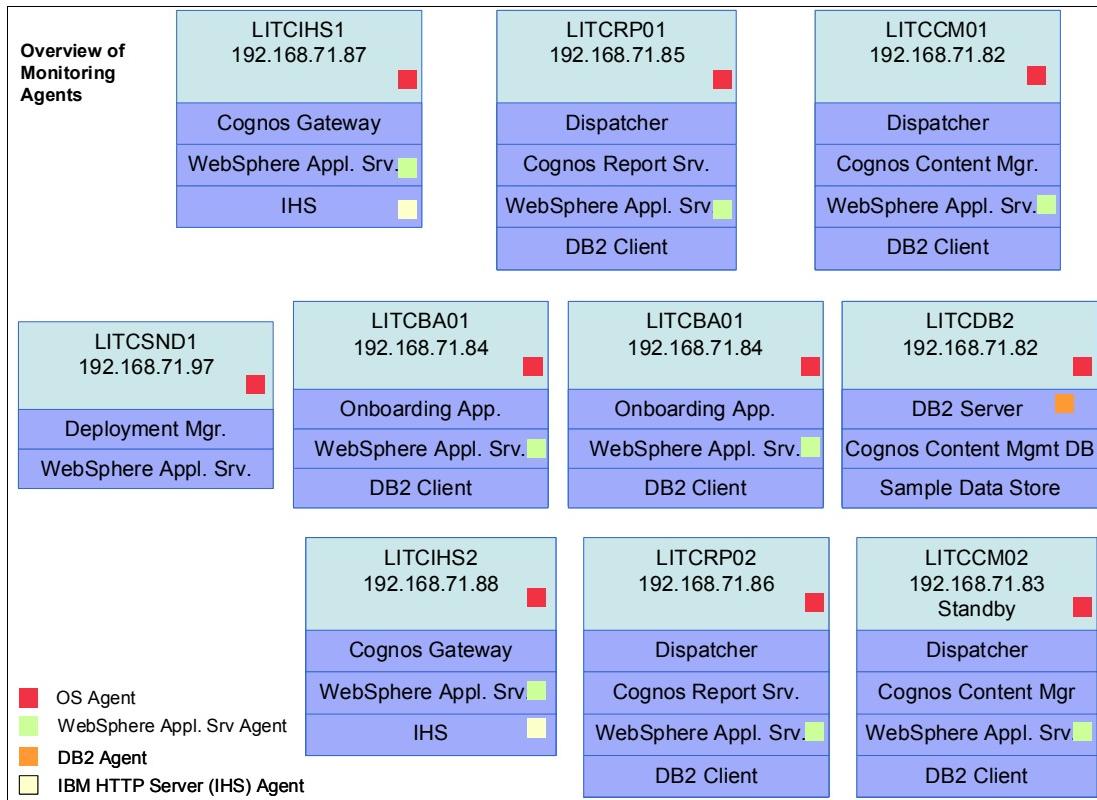

Figure 8-7   Tivoli Monitoring agents for the Smart Analytics Cloud

As already discussed and shown in Figure 8-6 on page 109, all agents connect to the central TEMS, and then the central TEPS is used to view all of the monitoring and event data centrally.

8.3 OSS: Provisioning management

The Smart Analytics Cloud can grow as needed by adding servers to expand capacity in production or creating additional systems for development teams. Automated provisioning of new virtual systems reduces the time the service administrators need to set up a new system to a great extent. This is true even for such flexible platform as System z. Adaptable and automated best practices for building and managing such an environment are applied to accelerate installation time.

8.3.1 Provisioning functions

In the context of the Smart Analytic Cloud, the important provisioning aspects that must be addressed are:

► Provision operating systems

► Provision middleware using cloning or installation and customization scripts

► Discover, track, and report data center resources to enable accurate server provisioning and software deployments

► Follow policies and preferred configurations in support of Cloud standards, corporate, and regulatory compliance efforts

► Automatically provision software and configurations to a wide variety of hardware and operating systems

► Maintain configurations and manage changes to resources

► Distribute software and patch management

► Automated deployment of servers through software templates

8.3.2 SAC provisioning solutions

When deciding about the implementation of a provisioning solution you have two alternatives, which we describe in the next two sections.

Tivoli Services Automation Manager

IBM Tivoli Services Automation Manager (TivSAM) is positioned as one of the strategic products for IBM cloud computing.

Tivoli Service Automation Manager (TivSAM) assists in the automated provisioning, management, and deprovisioning of hardware servers, networks, operating systems, middleware, and application-level software. Several

virtualization environments (hypervisors) are supported in the process of virtual server provisioning on IBM System x®, System p®, and System z.

TivSAM and the fully-integrated Tivoli Provisioning Manager provisioning product can be used to create and rapidly provision Linux on System z guests in the System Analytics Cloud. The hour or more of manual Linux install and configuration time can be automated and reduced to a matter of minutes.

TPM is the engine that runs work flows on behalf of the TivSAM product. The TivSAM user interface drives most of its functions through TPM work flows.

Figure 8-8 shows the topology of the Tivoli Service Automation Manager.
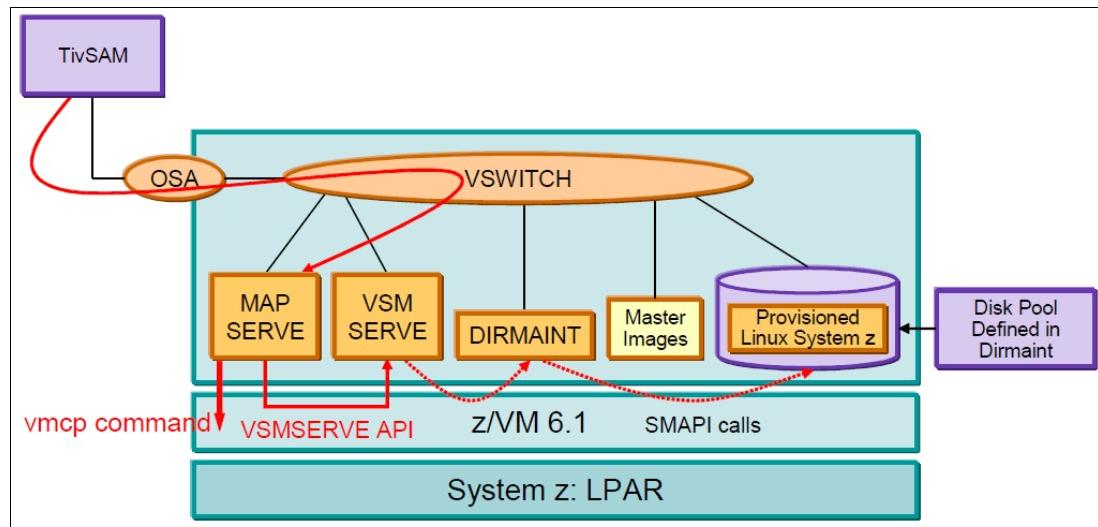
Figure 8-8   TivSAM topology

The network in this implementation is a VSWITCH. TPM work flows are run on the TivSAM server and the z/VM and Linux-specific provisioning steps are passed to the MAPSERVE Linux guest. MAPSERVE acts as a central point of control, executing z/VM commands directly or passing VM SMAPI commands to the VSMSERVE z/VM service machine through remote procedure calls (RPC). Some VM SMAPI commands generate DIRMAINT commands that are then passed to the DIRMAINT service machine. MAPSERVE uses the VMCP interface to run z/VM CP and CMS commands directly, usually under the MAPAUTH user ID. While a DIRMAINT disk pool is shown, dedicated disks can also be used when a Linux golden master image is created using a dedicated disk.

For more information about how to implement it, see *Provisioning Linux on IBM System z with Tivoli Service Automation Manager (TivSAM)*, REDP-4663.

IBM Mobility Workbench

The IBM Mobility Workbench is an architectural framework to support migration of distributed applications from a given source environment to a target environment. Fast multiplatform systems and software provisioning is one part of its functionality, which we used in our solution.

Figure 8-9 shows an overview of the IBM Mobility Workbench.
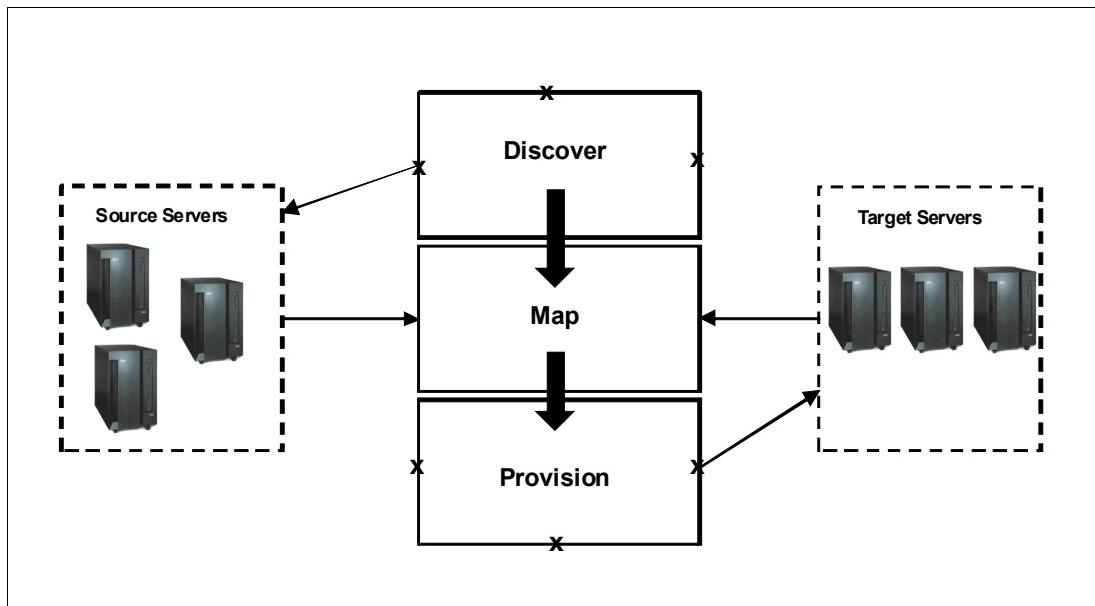
Figure 8-9   IBM Mobility Workbench structure overview

The IBM Mobility Workbench consists of several independent components that focus on three primary areas:

► Discover collects detailed information from the existing source servers and stores that information in a centralized and persistent location.

► Map (design) is responsible for mapping the source server configuration to the new target (to be) server configuration, which includes using the information from discovery and general information about the target environment.

► Provision builds the new server and migrates application configurations from the source to the target.

For the provisioning phase, an asset library is used to store the golden images (or golden clones) and the software binaries. At provisioning time, IBM Mobility Workbench retrieves requests from the database, provisions the Linux server using the golden clone, adds file systems and users groups, installs software products, and executes custom post-provisioning scripts.

8.4 BSS: Billing management

The values of cloud computing to the user and the enterprise means improving business performance and lowering the cost of delivering IT services to the organization. It also means acquiring computing services quickly without requiring understanding of the underlying technology and with visibility of their true IT cost. It delivers services for consumer and business needs in a way that is:

► Simple and easy to use
► Unbounded in scale
► Standardized
► Rapidly provisioned
► Billed accurately based on usage
► Changing consumption behavior
► Cost effective

8.4.1 Billing management functions

Billing Management considerations are fundamental to profitable cloud service delivery and cost transparency. Usage-based pricing in cloud delivery requires knowledge of service usage and resource usage. As organizations move to cloud, they must be able to demonstrate usage-based pricing. One of the infrastructure characteristics for effective cloud delivery is the rapid provisioning of server stack and other resources with seamless accounting for resource use.

What is needed to do resources usage and accounting:

► Know who is consuming which IT resources. Various data collectors for IT infrastructure are required to review the resources consumption across multiple dimensions.

► Determine the cost of those resources, including those that are shared in the cloud requires a sophisticated costing engine that assigns cost to resources usage.

► Apply and allocate cost for chargeback, ROI, costing analysis, and billing needs, costing, and reporting engine associates usage costs to consumers of IT resources.

Although resources and usage accounting component is one of the key building blocks, in this publication we only describe the IBM solution and products' capability on resources and usage accounting that we are not implementing in our lab environment. The resources and usage accounting is not used in our Smart Analytic Cloud project at this stage. It can be implemented next phase in the future to expand the Smart Analytic Cloud capability.

8.4.2 Billing with IBM Tivoli Usage and Accounting Manager

The IBM Tivoli Usage and Accounting Manager (TUAM) software family measures, collects, analyzes, reports, and, optionally, bills the utilization and costs of different computing resources including applications, servers, storage, networks, databases, messaging, and many other shared services in the cloud. TUAM consolidates all types of resources usage records into a common and integrated reporting structure with its rich Data Collectors associated with Operating Systems, databases, Internet infrastructure, network and printing, and customized usage data import collection from any application or system. TUAM then generates reports, invoices, and summarized files showing resource consumption and monetary expenditures for each functional unit within the organization.

The TUAM Data Collector collects information, d allocates costs to the appropriate user, and provides support for a wide range of z/OS and System z sources. TUAM also provides cross platform detailed reports and invoicing and enables measurement of resource usage and effective charge back. TUAM creates customizable web-based reports with an instant drill-down capability and lets you invoice users based on popular cost allocation methods.

TUAM allows alignment of IT costs with company priorities and can account for individual department's use of key applications, servers, and other IT resources by providing an extremely flexible end-to-end tool that helps businesses improve IT cost management. The TUAM business rule allows you to track, allocate, and invoice based on actual resource use by department, user, and many additional criteria.

Tivoli Usage and Accounting Manager components
This section describes the components of the IBM Tivoli Usage and Accounting Manager. Figure 8-10 on page 117 shows the high-level components of TUAM.
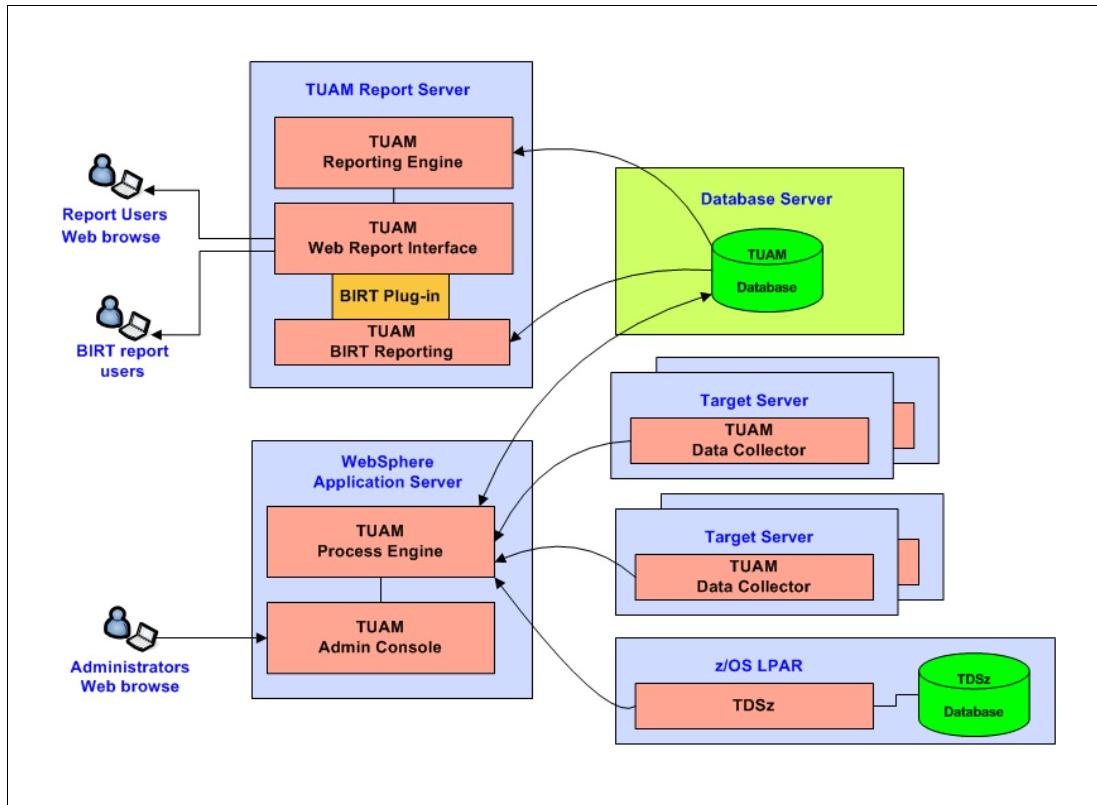
Figure 8-10   IBM Tivoli Usage and Accounting Manager

Data Collectors

TUAM provides powerful data collection agents and interfaces that process a wide variety of usage metering statistics. The collectors process across multiple platforms and support many different releases and formats of the various logs. It includes powerful summarization and account lookup features. TUAM Data Collectors provide a completely automated solution for accounting file management and integrate the data to a central server. The TUAM collection process can be fully automated either using TUAM's Job Runner utility or supported third-party scheduling systems. The collectors are not limited to machine metrics. In fact, TUAM can process virtually any type of input file including SMF, Machine Logs, spreadsheets, DBMS tables, Third Party Software Extracts, and so on.

Process Engine

TUAM data processing engine can process and apply business rules to large volumes of data collected from various sources and format. Data is stored on the

TUAM database. The TUAM database provides all of the data needed to produce complete and detailed reports, spreadsheets, and graphs.

Web Reporting

TUAM provides web-based reporting and graphing. Web reporting allows you to process, access, and analyze IT resource usage metrics from many IT systems and view the results right in your web browser. Reports, spreadsheets, and graphs can be generated and viewed through the Web Batch Report Generation and ad-hoc reporting.

Integrating with other Tivoli products

IBM Tivoli Usage and Accounting Manager integrates with IBM Tivoli Monitoring (ITM) and the Tivoli Data Warehouse (TDW) infrastructure as a new source to extract accounting metrics for Power Management and cloud architecture.

ITCAM for SOA discovers the services and collects consumption data and understand who is using the services. Data is feed to Tivoli Usage and Accounting Manager and delivers detailed information and reports about the intricate use of shared services so their usage can be optimized.

TUAM also integrates with IBM Tivoli Decision Support for z/OS, which provides a robust data collection facility for gathering usage and accounting data in the z/OS environment. Tivoli Usage and Accounting Manager for z/OS Option collects mainframe resource usage information from Tivoli Decision Support for z/OS and allocates costs to the appropriate user or department.

IBM Tivoli Service Automation Manager (TSAM) produces service usage data and integrates with Tivoli Usage and Accounting Manager (TUAM). It provides metering data to TUAM to deliver accurate resource usage information. IBM Tivoli Service Automation Manager for Linux on System z provides users with a self-service interface to select and reserve IT resources available in an IT service catalog. Through this self-service portal, end users can reserve resources needed to meet project requirements and schedule their deployment, thus eliminating lengthy, often error-prone and undocumented exchanges.

IBM Tivoli License Compliance Manager for z/OS can feed z/OS product and software use information to TUAM to help with cost allocation. Tivoli Asset Management for IT can provide assets costs to TUAM. TUAM provides web-based IT Financial Management reports that you can drill down.

Figure 8-11 on page 119 shows a sample TUAM integration with other Tivoli Service Management products.
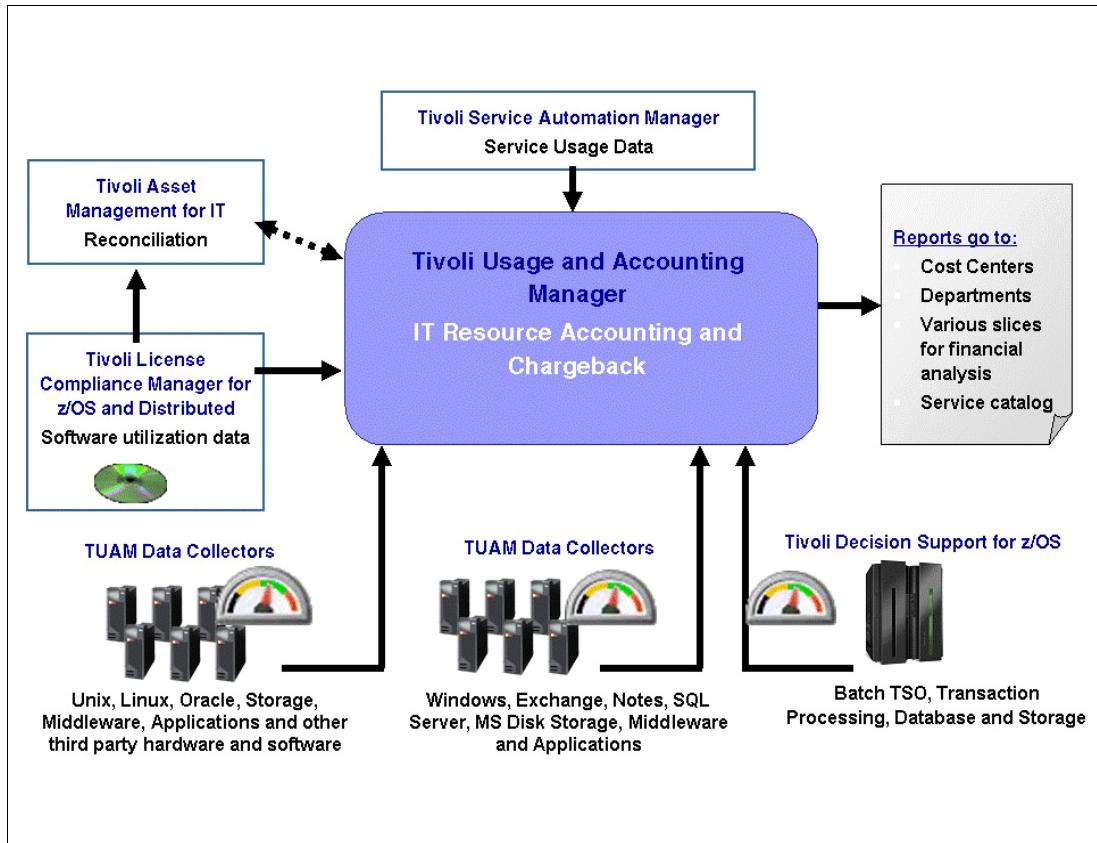
Figure 8-11   Tivoli Usage and Accounting Manager integrating with other Tivoli products

TUAM infrastructure for Smart Analytics Cloud

Although the resources and usage accounting component is one of the key building blocks, we only describe the IBM solution and products capability on resources and usage accounting in this publication. The resources and usage accounting is not used in our Smart Analytic Cloud project at this stage. It can be implemented next phase in the future to expand the Smart Analytic Cloud capability.

Figure 8-12 on page 120 shows the infrastructure that can be implemented for our lab environment for testing purposes in future. The Tivoli Usage and Accounting Manager Application Server component supports SLES11, and it can be implemented as a separate Linux on System z virtual server.

The Tivoli Usage and Accounting Manager database uses the existing DB2 database server. Because we have no z/OS in our environment, the TDS for

z/OS interface for TUAM is not required. We implemented the Report Server for Web Reporting on a Windows 2003 server.
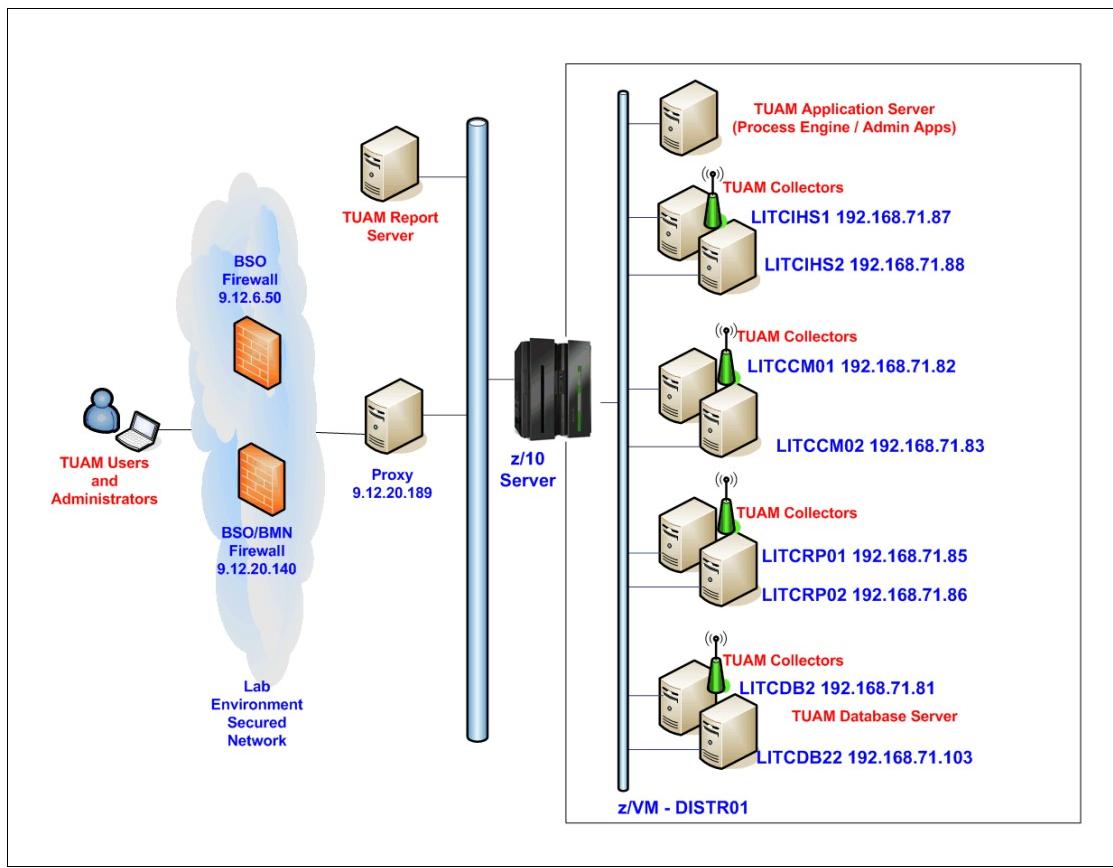

Figure 8-12   TUAM infrastructure for our lab environment

8.5 BSS: Ordering and offering management

This section focuses on the business support services managing orders and offerings.

8.5.1 SAC ordering and offering functions

In the context of the Smart Analytic Cloud, the important ordering and offering functions that must be addressed are:

► Provide an onboarding interface for end-users to create a request and collect information about the Cloud service request. We implemented that in our lab environment by the onboarding application (see section 6.2, "Onboarding application" on page 66 and Chapter 12, "Onboarding application" on page 211).

► Manage catalog-based requests.

► Provide workflow capabilities for ordering and requests management.

► Allow searchable solutions to give quick access to solutions for specific service requests.

► Publish Cloud service offerings to give users information about what services are available to them.

► Provide cost information for service to associate cost to the cloud service offerings, which helps users understand costs and manage consumption.

► Provide service entitlement so that service catalog entries can be provided based on a group or business unit.

► Provide Service Definition Templates to provide templates for common service items.

8.5.2 SAC solution with Tivoli Service Request Manager

Built from the ground up on the Information Technology Infrastructure Library (ITIL®) framework, Tivoli Service Request Manager® provides a comprehensive and modular approach to integrated Service Desk and Service Catalog management. In addition, using this product you can establish and efficiently operate a corporate service desk for service requests around enterprise assets. Tivoli Service Request Manager runs on top of Tivoli's process automation engine (Tivoli Automation Manager - TAM), utilizing its services and J2EE technology with advanced business process management (based on SOA, web services, and XML). This allows adaptation, configuration, rapid deployment, robust scalability, and deep integrations with enterprise applications.

Tivoli Service Request Manager is introduced as a common package of Service Desk and Service Catalog capabilities, as shown in Figure 8-13 on page 122. What Service Desk and Service Catalog have in common is the concept of satisfying user requests. In general, requests to the Service Desk are handled on a case-by-case basis, while Service Catalog requests are usually handled in a

pre-configured manner that is intended to fulfill the request in a best practice manner.
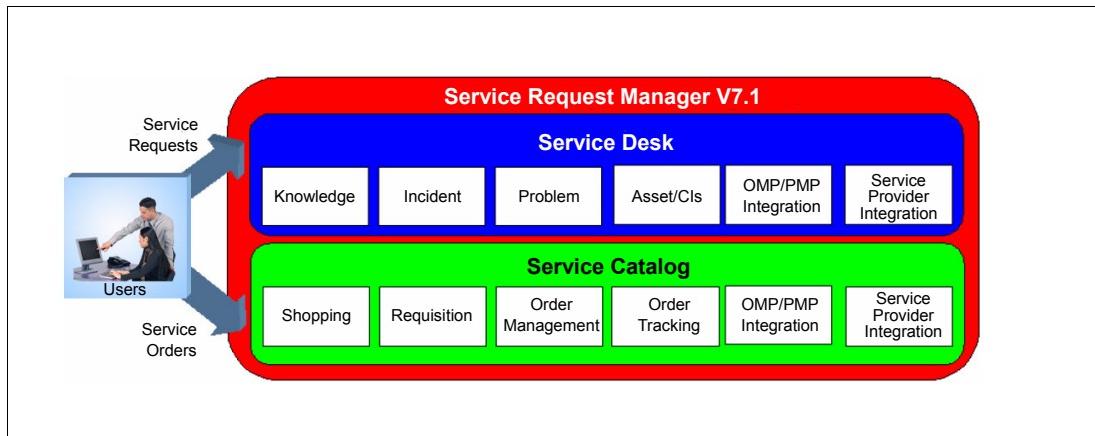
Figure 8-13   Tivoli Service Request Manager structure

The Service Catalog supports and facilitates the processes defined by customers by addressing the issues discussed.

Service Catalog has three types of supported services:

► Descriptive services: These services permit a company to advertise services that are delivered in ad hoc, manual manners. It allows a customer to put a service in the Service Catalog in a low-cost manner.

► Action: Action services provide a cheap way to integrate existing automated services through Launch in Context to external applications or by the execution of Tivoli Service Request Manager Actions, which can do things, such as, trigger command scripts or custom Java classes that call external APIs.

► Supply Chain: The Supply Chain service fulfillment model leverages the full power of the Tivoli Service Request Manager purchasing supply chain applications. Service Catalog Supply Chain refers to the chain of Service Catalog product components that accomplishes the full set of Service Delivery tasks. The chain starts with the ability of a Service Requester to search for and requisition a service from the Service Catalog. The chain ends with the complete fulfillment of the Service Requisition.

Tivoli Service Request Manager provides a detailed understanding of the Cloud services delivery:

► Which Cloud services are actually being provided

► How frequently they are being requested

- How well they are being fulfilled
- Allocate existing resources to higher priority tasks and prevents incidents with end-user self-service
- Improves quality and speed of service delivery—automated service fulfillment process is repeatable and consistent
- Speed and flexibility—internal customers can shop for published IT service offerings through service catalog 24x7
- Configure the visibility to service offerings based on location and department
- Measure end-user satisfaction through built-in surveys and track the trends
- Measure SLA compliance to service requests

8.6 Security enforcement and management

This section focuses on the operational support services that enforce and manage security. To ensure that all aspects are addressed we use the IBM security architecture blueprint as a reference architecture. The blueprint covers three areas:

- Business Security References Mode
- Foundational Security Management
- Security Services and Infrastructure

For a detailed description of the IBM Security Architecture Blueprint, see *Introducing the IBM Security Framework and IBM Security Blueprint to Realize Business-Driven Security*, REDP-4528.

However, for our lab environment, we refer to a subset of this blueprint, which is explained in further detail in 8.6.2, "Cloud security" on page 125.

8.6.1 IBM Security Architecture Blueprint

The IBM Security Blueprint uses a product-agnostic and solution-agnostic approach to categorize and define security capabilities and services that are required to answer business security requirements or areas of concern categorized by the IBM Security Framework.

In the blueprint, IBM aims to identify architectural principles that are valid across all domains and fundamental services within and across the domains. It also highlights applicable best practices and IT standards.

The blueprint was created based on researching many customer-related scenarios focusing on how to build IT solutions based on the IBM Security Framework. The intention of the blueprint is that it can be used as a roadmap to assist in designing and deploying security solutions in any IT organization.

Figure 8-14 shows the IBM Security Architecture Blueprint.
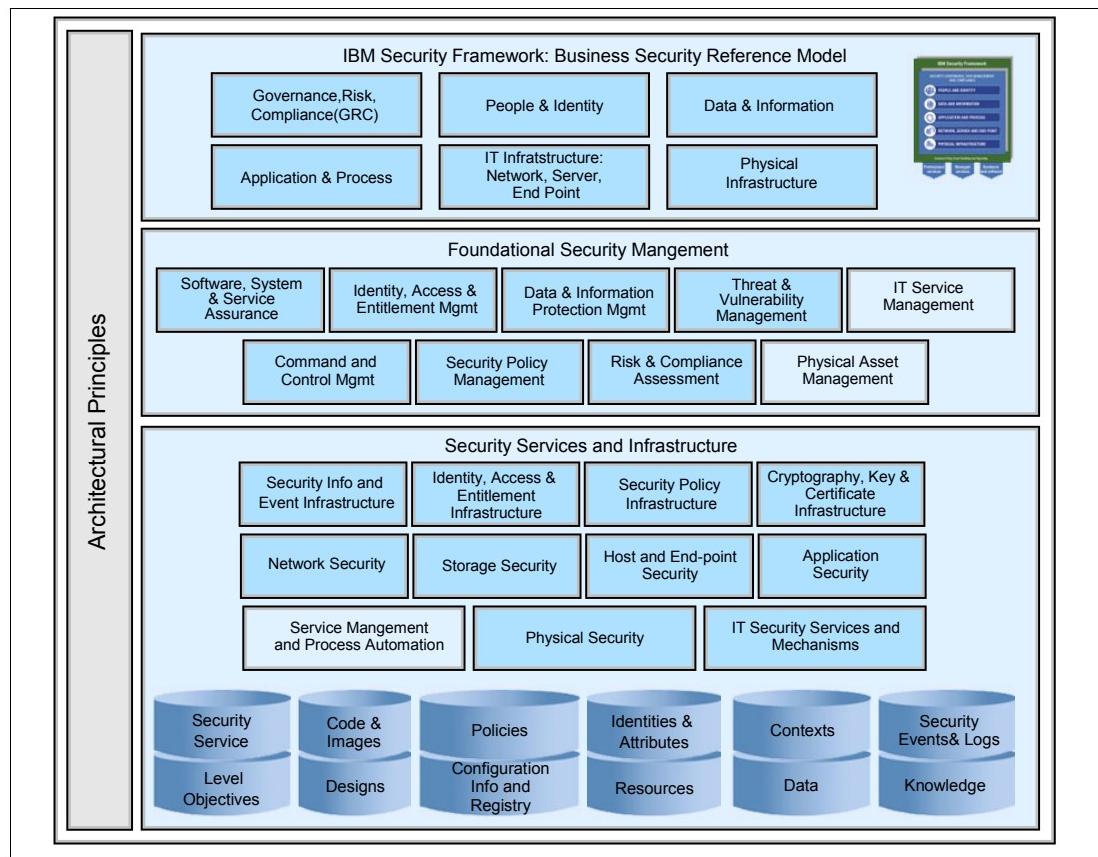
Figure 8-14   IBM Security Architecture Blueprint

This blueprint also applies to the Smart Analytics Cloud. The massive sharing of infrastructure with cloud computing creates a significant higher risk for cloud security compared to security in more traditional IT environments. Users spanning different divisions and trust levels often interact with the same set of computing resources. At the same time, workload balancing, changing service level agreements, and other aspects of today's dynamic IT environments create even more opportunities for misconfiguration, data compromise, and malicious conduct. Infrastructure sharing calls for a high degree of standardized and process automation, which can help improve security by eliminating the risk of

operator error and oversight. However, the risks inherent with a massively shared infrastructure mean that cloud computing models must still place a strong emphasis on isolation, identity, and compliance.

8.6.2 Cloud security

There is no one-size-fits-all model for security in the cloud. Organizations have different security requirements that are determined by the unique characteristics of the business workload that they intend to migrate to the cloud.

With Software as a Service (SaaS), most of the responsibility for security management lies with the cloud provider. SaaS provides a number of ways to control access to the web portal, such as managing user identities, application level configuration, and restricting access to specific IP address ranges or geographies.

With Platform as a Service (PaaS), clients assume more responsibilities for managing the configuration and security for the middleware, database software, and application runtime environments.

Infrastructure as a Service (IaaS) transfers even more control and responsibility for security from the cloud provider to the client. In this model, access is available to the operating system that supports virtual images, networking, and storage.

The following security measures represent general best practice implementations for cloud security. At the same time, they are not intended to be interpreted as a guarantee of success:

► Implement and maintain a security program.
► Build and maintain a secure cloud infrastructure.
► Ensure confidential data protection.
► Implement strong access and identity management.
► Establish application and environment provisioning.
► Implement a governance and audit management program.
► Implement a vulnerability and intrusion management program.
► Maintain environment testing and validation.

Figure 8-15 on page 126 shows the layers of the cloud architecture overview, as shown in Figure 8-1 on page 100, and the related security considerations.
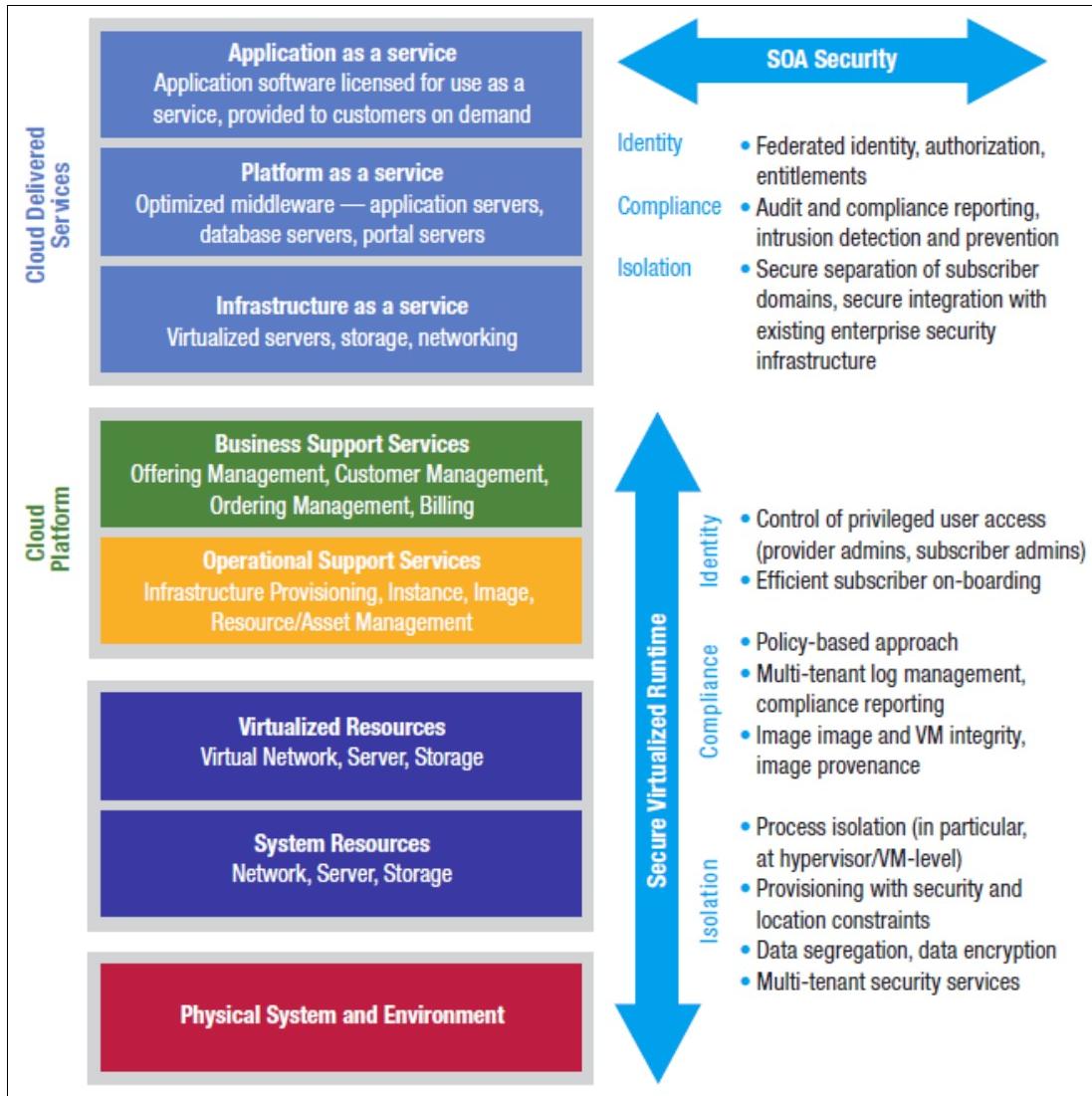
Figure 8-15   Layers of a typical cloud service and security

The cloud architecture that we introduced allows you to construct a model of cloud security consisting of two main concepts: an SOA security layer that resides on top of a Secure Virtualized Runtime layer.

The SOA Security Model, which is a subset of the IBM Security Architecture Blueprint, fully applies to the cloud because the cloud services has the characteristics of a distributed SOA environment. Different services and cloud

environments can be combined for new cloud environments within an enterprise. The services might be in different administrative or security domains that connect together to form a single cloud application. Further information about the SOA security model can be found in the IBM Redbooks *Understanding SOA Security*, SG24-7310-01. A basic concept in SOA is to externalize security into services and to make them available for use by other services.

The Secure Virtualized Runtime layer on the bottom of Figure 8-15 on page 126 is a virtualized system that runs the processes that provide access to data on the data stores. This run time differs from classic runtime systems in the way that it operates on virtual machine images rather than on individual applications. It provides security services, such as isolation, data segregation, control of privileged user access, provisioning with security and location constraints, image provenance, image and virtual machines integrity, multi-tenant security services, and intrusion prevention.

8.6.3 SAC security functions

In the context of the Smart Analytic Cloud, the important security concerns that must be addressed are:

► Web access control: Because the SAC user interface is web-based, the common risks that are related to internal or external web-applications do exists. These risks are higher in a multi-tenant and various user origins situation.

The web access control must provide the following security functions:

– Implement a least privilege model

– Centralized external authentication gateway

– Centralized external authorization provider

– Single Sign On with back-end application servers

– Rules based authorization engine

– Implement federated identity management

– Security must be externalized from the resource managers and centralized in the security manager

– Extended auditing and reporting capabilities

► Virtual resources security: Because the SAC is running on a virtualized infrastructure that can potentially be massively multi-tenant, isolation and segregation of all resources is of high importance.

The virtual resources security must provide the following functions:

– Isolation of virtual environments

– Isolation of virtual resources
  – Multi-tenant intrusion prevention
  – Extended auditing and reporting capabilities
  – Location constraints
  – Image provenance and image and virtual machines integrity
  – Multi-tenant security services

► Data security: Because the SAC accesses sensitive business information from many users from different origins or with various privileges, it is critical to enforce tight security controls on data access.

The data security must provide the following functions:

► Ensure confidential data protection
► Protect confidential and business critical data
► Protect application information
► Secure data communications
► Protect Personally Identifiable Information (PII)
► Securely destroy all non-essential PII
► Protect intellectual property
► Protect encryption keys from misuse or disclosure
► Implement data loss prevention
► Audit and report on sensitive data access

8.6.4 SAC web access control: Tivoli Access Manager for e-business

The Cognos user interface is web-based and requires adequate access control mechanisms. The core of the web access control requirements can be satisfied by the IBM Tivoli Access Manager for e-business software. This access management software from Tivoli addresses the difficulty of executing security policies over a wide range of web and application resources, such as in a Cloud. It supports authentication, authorization, data security, and resource management capabilities.

Figure 8-16 on page 129 shows a typical layout of the Tivoli Access Manager components for Internet security.
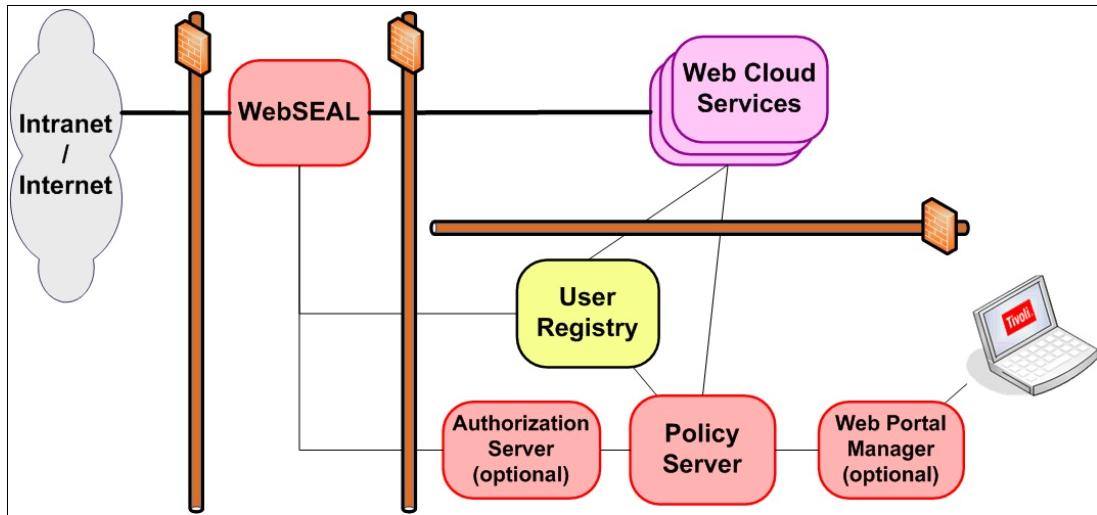
Figure 8-16   Typical Tivoli Access Manager components for Internet security

In a typical web infrastructure, application layer security is enforced using components that provide at least authentication services and authorization services. It is a best practice to place the authentication service in the DMZ so that end-user are filtered and known as early as possible. The authentication service is commonly provided by a Reverse Proxy Security Server (RPSS) in the DMZ with assistance from a User Registry for identities and a Security Manager for policies in the management zone. The authorization service can be provided by a Security Manager that possesses Access Control Lists (ACL) and policies. The security manager is responsible for the authorization decision-making process that helps to enforce security policies. Authorization decisions made by the authorization service result in the approval or denial of client requests to perform operations on protected resources.

There are advantages in externalizing the authentication and authorization services, for example, it centralizes the security management across heterogeneous environments. It also increases security by providing consistent and homogeneous role-based and policy-based security management.

A Tivoli Access Manager for e-business solution is usually composed of the following components:

► User registry: It supports the Access Manager authorization functions. The registry provides a database of the user identities known to Tivoli Access Manager. It also provides a representation of groups in Tivoli Access Manager roles that can be associated with users. Finally it provides a data store of metadata required to support additional functions.

- Policy Server: It maintains the master authorization database for the secure domain. This server is key to the processing of access control, authentication, and authorization requests. It also updates authorization database replicas using push/pull methods and maintains location information about other Tivoli Access Manager servers in the secure domain. There can be only one instance of the policy server and its master authorization database in any secure domain at one time. For availability purposes, a standby server can be configured to take over policy server functions in the event of a system failure.

- WebSEAL: It is a high-performance, multi-threaded reverse proxy security server (RPSS) that sits in front of back-end Internet applications. It applies a security policy to a protected object space. WebSEAL can provide single sign-on solutions and incorporate back-end Internet application server resources into its security policy. Because it is implemented on an HTTP server foundation, it is limited to enforcing policy for applications communicating with HTTP and HTTPS protocols. It uses junctions to define the connectivity to the multiple Web Cloud Services back-ends. It can propagate identities for Single Sign On.

- Authorization Server (optional): It can be installed to offload authorization decisions from the Policy Server and provide for higher availability of authorization functions. The Policy Server provides updates for authorization database replicas maintained on each Authorization Server. It is an optional component.

- Web Portal Manager (optional): It is an Internet-based graphical user interface (GUI) used for Tivoli Access Manager administration. Similar to the pdadmin command line interface, this GUI provides management of users, groups, roles, permissions, policies, and other Tivoli Access Manager tasks. A key advantage of the Web Portal Manager over the pdadmin command line utility is the fact that it is a browser-based application that can be accessed without installing any Access Manager-specific client components on the administrator's local machine or requiring special network configuration to permit remote administrator access.

For Single Sign On from the end-user workstation, software, such as IBM Tivoli Access Manager for Enterprise Single Sign-on, can be used.

For identity propagation, transformation, and mapping, software, such as IBM Tivoli Federated Identity Manager can be used.

For identity life cycle management and access rules provisioning, software, such as IBM Tivoli Identity Manager, can be used.

8.6.5 SAC virtual resources security

In the context of the Smart Analytics Cloud (SAC), the virtual resources are all that the infrastructure resource made available for Cognos to run. These resources include virtual CPU, virtual memory, virtual network, virtual machines for Linux, Java virtual machine, and so on.

Hardening

The first step in securing these is to apply the standard hardening techniques for each of the resource types and all layers, for example, the SLES Hardening Guide is available to help hardening and securing of the SAC Linux guests.

Refer to each software component documentation or manual for security hardening best practices.

Centralized security management

The second step is to apply the security policy and principles along with the best practices for the security management of these components. One important security principle is to externalize and centralize security management.

For the SAC Linux guests, it can be done for the operating system with software, such as IBM Tivoli Access Manager for Operating System. It protects individual application and operating system resources by addressing system vulnerabilities surrounding privileged user access (super user or root accounts). It offers centralized, policy-based user access management, tracking, and control in a heterogeneous OS environment.

Auditing and reporting

The third step is to increase the control by implementing proper auditing and reporting on all activities within the SAC.

Each component usually has its own auditing capabilities. The goal is to centralize the auditing capabilities and then extract some reports from the security logs and use them for compliance purposes, for example, at the SAC Linux level, if IBM Tivoli Access Manager for Operating System is used, the auditing capabilities of this product allows centralizing of all security log activities within one repository.

For properly managing and addressing information retrieved from or traversing all components in the SAC, software, such as IBM Tivoli Security Operations Manager (TSOM), can be used. It handles the analysis and correlation of security information from network security devices, access or identity management, and enterprise security applications.

Figure 8-17 shows the analysis and correlation that TSOM does.
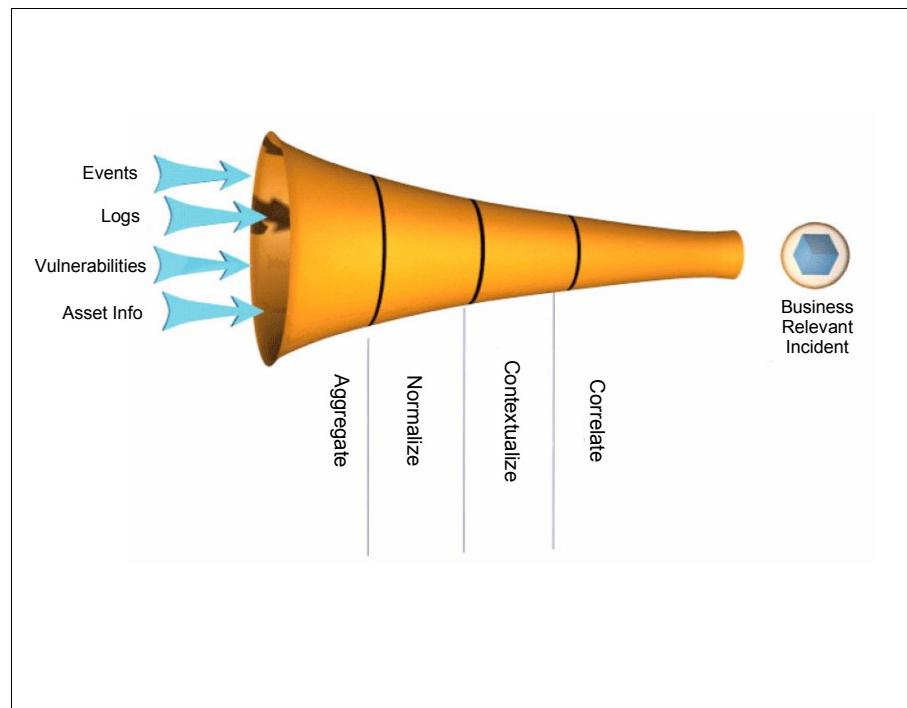
Figure 8-17   Overall view of analysis and correlation processes

8.6.6 SAC data security

In the context of the Smart Analytics Cloud, the data refers to not only the sensitive business data accessed by Cognos but also any data stored and used by the Cloud, such as end-user information or cloud services description or usage.

Hardening

The first step in securing these is to apply the standard hardening techniques for each of the data store types, for example, the DB2 UDB security best practices are available to help hardening and securing the SAC Cognos databases and the onboarding application and LDAP databases.

Refer to each data store component documentation or manual for security hardening best practices.

Granular access control

In any type of multi-tenant infrastructure, such as a cloud, it is important to assign granular identities to requests so that granular access controls can be made. The ideal case scenario is to have the end-user identity being propagated from the external authentication gateway, to the application server and to the database. This way data authorization based on user identity or group membership can be enforced and granular data access auditing can be executed.

The techniques and mechanisms for identity propagation are often product specific. In the context of the SAC, Tivoli Access Manager WebSeal authentication gateway can forward the identity to Cognos running on top of WebSphere Application Server. There are ways to configure the data source between WebSphere and DB2 so that the end user identity gets propagated. Also, products, such as Tivoli Federated Identity Manager, can be used if mapping is necessary.

Data segregation

For data segregation issues and managing shared data resources within a multi-tenant environment, the IBM Systems, Storage, and Network Segmentation Solutions offer application isolation, OS containers, encrypted storage, VLANs, and other isolation technologies for a secure multi-tenant infrastructure. Designed to be shared by thousands of users, the IBM server has security built into nearly every level of the computer from the processor to the OS to the application level.

Encryption

Encryption is the most common way to enforce data confidentiality.

In the context of the SAC where business data is stored in a DB2 database, IBM Database Encryption Expert can be used. It enables encryption of data so that it is possible to have data in the cloud but be able to control who can access it. It can protect sensitive information in both online and offline environments and has centralized policy and key management to simplify data security management.

For all other data stored and used by the cloud, refer to each product documentation to know more about the encryption capabilities.

Encryption of data in flight can be enforced between the end-user and Cognos using HTTP over SSL (HTTPS), which is a standard feature provided by the underlying WebSphere Application Server.

Auditing and reporting

Each data store component usually has its own auditing capabilities. The goal is to centralize the auditing capabilities and then extract some reports from the security logs and use them for compliance purposes.

As described earlier, Tivoli Security Operations Manager (TSOM) can be used.

Part 4

Implementation

In this part, we discuss how we implemented a Smart Analytics Cloud in our lab environment.

9

WebSphere infrastructure for Cognos 8 BI

The IBM WebSphere Application Server is the foundation of the IBM WebSphere software platform and a key building block of a service-oriented architecture (SOA). It is available in various packages that are each designed to meet a certain type of client requirement. The core in each package is a WebSphere Application Server that provides the runtime environment for enterprise applications and an application server. All IBM WebSphere Application Server packages support a single stand-alone server environment and a distributed server environment. The benefits of WebSphere Application Server to the enterprise include:

► Simple, rapid development and deployment
► Secure, scalable, and highly-available run time
► Extensive communication services
► Effective security and application management

Because of these advantages, we choose to run IBM Cognos 8 BI on the IBM WebSphere Application Server. IBM WebSphere Application Server provides a secure and scalable application infrastructure for the IBM Cognos 8 BI Service Oriented Architecture.

137

9.1 WebSphere infrastructure for IBM Cognos 8 BI

The IBM WebSphere Application Server runtime environment was set up for the Cognos 8 BI components. The default application server shipped with Cognos 8 BI and used during the Cognos installation is Apache Tomcat, which is an open source software implementation of the Java Servlet and JavaServer Pages technologies. Cognos supports IBM WebSphere Application Server, but a stand-alone server profile is used as the default installation. To take advantage of the benefits and values of the IBM WebSphere Application Server Network Manager or an existing WebSphere Application Server runtime environment, we decided to install the Cognos 8 BI components on an IBM WebSphere Application Server Network Deployment runtime environment. To support high availability and scalability, we set up clustered servers for Cognos applications. We also integrated the WebSphere Application Server environment with an external LDAP server for user authentication.

In addition, we leveraged the strengths of the IBM z10 server and Linux on System z virtualization technology for building the IBM WebSphere Application Server infrastructure for Cognos 8 BI. All of the Cognos 8 BI components are installed and running on their own Linux on System z virtual server.

Figure 9-1 on page 139 shows the architecture for the WebSphere infrastructure for Cognos 8 BI components.

All Linux System z servers run as a z/VM guest on a single z/VM LPAR. Except for the Deployment Manager server, all other application servers are clustered under WebSphere Application Server for the Cognos application. There are three IBM WebSphere Application Server clusters created and prepared for the subsequent Cognos 8 BI installation. The LDAP server for user authentication for the IBM WebSphere Application Server and Cognos is outside our lab environment. The database component of the Cognos 8 BI is not shown in Figure 9-1 on page 139 for simplicity but is depicted in Chapter 10, "Cognos installation" on page 173.

The fourth and the fifth application clusters are created and used for our own LDAP web admin application and our test onboarding application, neither of them is part of Cognos 8 BI.

The ITSO firewalls and the ITSO Proxy server demonstrate a typical set up of a real client installation that separates the network zones at various security levels.
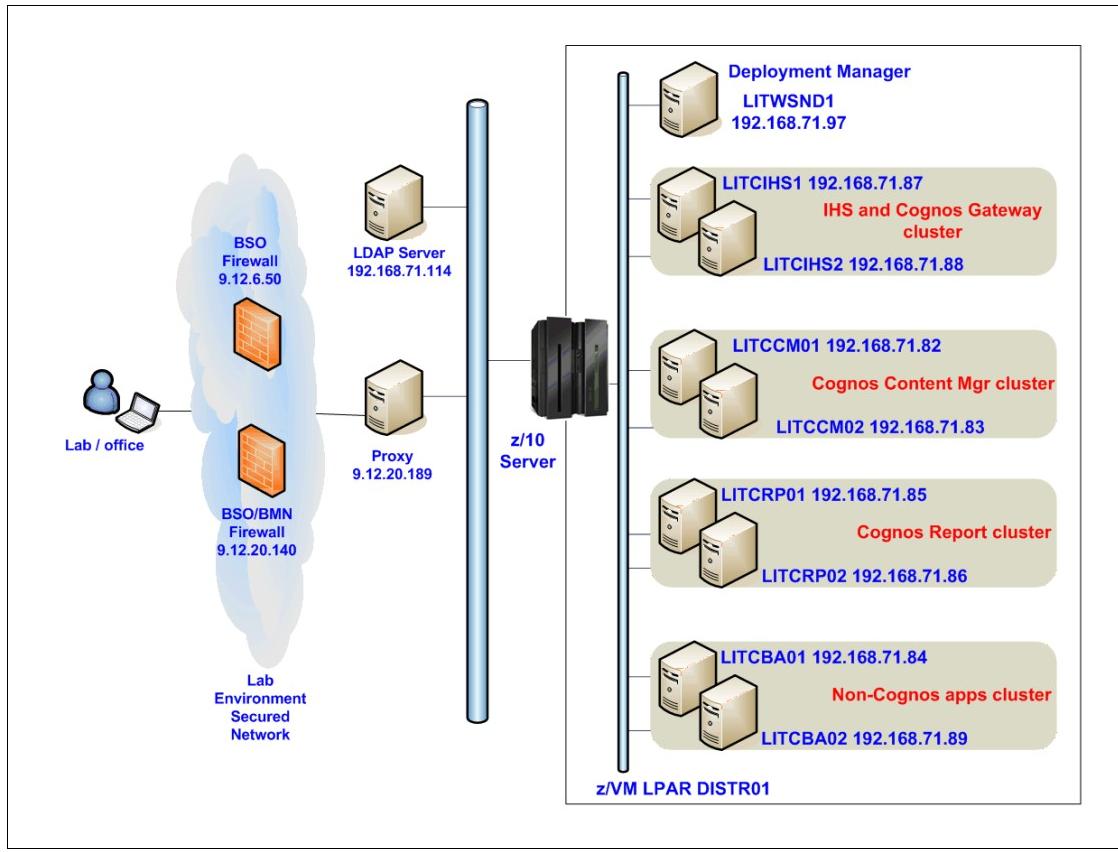
Figure 9-1   WebSphere Infrastructure for Cognos 8 BI

As shown in Figure 9-1, four pairs of Linux guests are used for the required application clusters. Three of them are used for the Cognos 8 BI application and the fourth one is used for the non-Cognos applications:

► A Cognos Gateway cluster for Cognos 8 BI Gateway Server

► A Cognos Report cluster for Cognos 8 BI Report Server

► A Cognos Content Manager cluster for Cognos 8 BI Content Manager Server

► Two non-Cognos application clusters for the LDAP web client and our onboarding application

9.2 Our lab environment

In this section, we describe some high-level system requirements of WebSphere Application Server for setting up the Cognos 8 BI infrastructure. Novell SUSE Linux Enterprise Server (SLES) was used as the operating system. IBM WebSphere Application Server V7.0 for Linux on System z was used in our lab environment as the target application server for Cognos 8 BI. All our SLES servers were running as Linux guests under one z/VM LPAR. Listed here are the software products that are used for the Cognos 8 BI V8 installation:

► IBM WebSphere Application Server Network Deployment v7.0 for Linux on System z 64-bit

► IBM HTTP Server V7.0 and WebSphere Application Server Plug-in

► IBM WebSphere Application Server V7 Fix Pack 9

► SDK Fix Pack 9 for IBM WebSphere Application Server V7

Prior to installing WebSphere Application server, the following system components must be ready:

► IBM z10 Server
► z/VM v6.1 LPAR
► SLES11 64-bit guests
► IBM Tivoli Directory Server v6

9.3 Installing the IBM WebSphere Application Server components

In this section, we give a high-level description of how the IBM WebSphere Application Server is installed. Nine virtual SLES11 Linux guests were created ahead of time and used for the IBM WebSphere Application Server installation. They serve as the foundation and the target runtime environment for Cognos 8 BI.

Table 9-1 gives a list of the Linux guests with their IP addresses and a description of their function.

Table 9-1   SLES11 Linux guests for IBM WebSphere Application Server installation

Host name	IP address	Descriptions
LITWSND1	192.168.71.97	Deployment Manager Server
LITCIHS1	192.168.71.87	HTTP Server for Cognos Gateway Server

Host name	IP address	Descriptions
LITCIHS2	192.168.71.88	HTTP Server for Cognos Gateway Server
LITCBA01	192.168.71.84	Application server for non-Cognos applications
LITCBA02	192.168.71.84	Application server for non-Cognos applications
LITCCM01	192.168.71.82	Application server for Cognos Content Manager
LITCCM02	192.168.71.83	Application server for Cognos Content Manager
LITCRP01	192.168.71.85	Application server for Cognos Report Server
LITCRP02	192.168.71.86	Application server for Cognos Report Server

Table 9-2 shows the software package that was used for IBM WebSphere Application Server V7 installation.

Table 9-2   IBM WebSphere Application Server installation packages

Component	Package name
WebSphere Application Server ND v7.0	WebSphere Application Server ND 64-bits for Linux on z - C1G3TML.tar.gz
WebSphere Application Server v7 Update Installer	WebSphere Application Server Supplements 1 - C1G3PML.tar.gz
WebSphere Application Server v7 Fix Pack 9	7.0.0-WS-WebSphere Application Server-LinuxS39064-FP0000009.pak
WebSphere Application Server SDK Fix Pack 9	7.0.0-WS-WASSDK-LinuxS39064-FP0000009.pak

For the installation, the shared disks, that store the installation binaries, are NFS mounted to the target SLES11 Linux guests, as shown in Example 9-1.

Example 9-1   Sample mount command and display the mount point after the mount

```
mount 192.168.71.108:/pub/itso /mnt

df - h
Filesystem              Size  Used Avail Use% Mounted on
192.168.71.108:/pub/itso
                        325G  285G   24G  93% /mnt

ls -al /mnt
drwxr-xr-x  3 1316  764 4096 Apr  8 16:09 WAS7
```

9.3.1 Installing the IBM WebSphere Application Server

To allow an installation that is easily repeated and that the provisioning tool can automate, we used the Silent Install Mode to install IBM WebSphere Application Server V7. To install the bases:

1. Expand the IBM WebSphere Application Server V7 package tar file.

2. Go to the WebSphere Application Server directory, and modify the responsefile.nd.txt file, as demonstrated in Example 9-2.

Example 9-2   Sample responsefile.nd.txt

```
-OPT silentInstallLicenseAcceptance.value="true"
-OPT disableOSPrereqChecking="true"
-OPT installType="installNew"
-OPT profileType="none"
-OPT installLocation="/opt/IBM/WebSphere/AppServer"
```

3. Start the silent install using the command in Example 9-3.

Example 9-3   Sample install command

```
./install -options "responsefile.nd.txt" -silent
```

4. Check the installation log and verify the install result. Example 9-4 shows a sample of our installation results.

Example 9-4   Sample install log and result

```
tail -2 /opt/IBM/WebSphere/AppServer/logs/install/log.txt

(Apr 1, 2010 3:15:01 PM), Process,
com.ibm.ws.install.ni.ismp.actions.SetExitCodeAction, msg1, CWUPI0000I:
EXITCODE=0
(Apr 1, 2010 3:15:01 PM), Process,
com.ibm.ws.install.ni.ismp.actions.ISMPLogSuccessMessageAction, msg1,
INSTCONFSUCCESS
```

9.3.2 Installing the UpdateInstaller and fix pack for WebSphere Application Server

After the base is installed successfully, continue to install the UpdateInstaller and the fix pack for the IBM WebSphere Application Server:

1. Expand the UpdateInstaller package tar file.

2. Go to the UpdateInstaller directory, and modify the
   responsefile.updiinstaller.txt file, as shown in Example 9-5.

Example 9-5   Sample responsfile.updiinstaller.txt file

```
-OPT silentInstallLicenseAcceptance="true"
-OPT disableOSPrereqChecking="true"
-OPT installLocation="/opt/IBM/WebSphere/UpdateInstaller"
```

3. Start the silent install for the UpdateInstaller, as shown in Example 9-6.

Example 9-6   Sample UpdateInstaller install command

```
./install -options "responsefile.updiinstaller.txt" -silent
```

4. Check the installation log, and verify the install result. Example 9-7 shows the
   result of a sample install log.

Example 9-7   Sample UpdateInstaller install log and result

```
tail -3 /opt/IBM/WebSphere/UpdateInstaller/logs/install/log.txt

(Apr 1, 2010 5:51:53 PM), Install,
com.ibm.ws.install.ni.ismp.actions.SettleNIFRegistryAction, msg1,
Current install/uninstall process is successful. Process type is:
install
(Apr 1, 2010 5:51:53 PM), Install,
com.ibm.ws.install.ni.ismp.actions.SetExitCodeAction, msg1, CWUPI0000I:
EXITCODE=0
(Apr 1, 2010 5:51:53 PM), Install,
com.ibm.ws.install.ni.ismp.actions.ISMPLogSuccessMessageAction, msg1,
INSTCONFSUCCESS
```

5. After the UpdateInstaller is installed successfully, go to the
   /opt/IBM/WebSphere/UpdateInstaller/responsefiles directory.

6. Modify the file install.txt for installing the fix pack for WebSphere Application
   Server, as shown in Example 9-8.

> **-W option:** Enter the -W option on one single line.

*Example 9-8   Sample install.txt file for installing fix pack for WebSphere Application
Server*

```
-W
maintenance.package="/mnt/WAS7/7.0.0-WS-WAS-LinuxS39064-FP0000009.pak"
```

```
-W product.location="/opt/IBM/WebSphere/AppServer"
```

7. Run the **update** command to install the fix pack, as shown in Example 9-9.

> The **update command:** Enter the **update** command on one single line.

Example 9-9   Sample update command with -silent option

```
./update.sh -options
"/opt/IBM/WebSphere/UpdateInstaller/responsefiles/install.txt" -silent
```

8. Check the installation log, and verify the install result.

9. Repeat the steps from updating the install.txt file to install the WebSphere Application Server SDK Fix Pack.

> **TIP:** After the IBM WebSphere Application Server base and fix packs were installed, we cloned the disk with /opt/IBM/WebSphere to all other Linux guests using flashcopy instead of repeating the installation steps on all other Linux guests, and proceeded to create the server profiles.

9.3.3 Installing the IBM HTTP Server

We also used the Silent Install Mode to install IBM HTTP Server (IHS) V7 on Linux guest LITCIHS1, LITCIHS2, and LITCCM01. Table 9-3 shows the software packages that we used for the IBM HTTP Server V7 installation.

Table 9-3   IBM HTTP Server installation packages

Component	Package name
IBM HTTP Server	WebSphere Application Server Supplements 1 - C1G3PML.tar.gz
UpdateInstaller	WebSphere Application Server Supplements 1 - C1G3PML.tar.gz
IBM HTTP Server Fix Pack	7.0.0-WS-IHS-LinuxS39064-FP0000009.pak

To install the IBM HTTP Server base:

1. Expand the IBM HTTP Server package tar file.

2. Go to the IHS directory, and modify the responsefile.txt file, as shown in Figure 9-10 on page 145.

Example 9-10   Sample responsefile.txt for IHS Server install

```
-OPT silentInstallLicenseAcceptance="true"
-OPT disableOSPrereqChecking="true"
-OPT installLocation="/opt/IBM/HTTPServer"
-OPT httpPort="80"
-OPT adminPort="8008"
-OPT createAdminAuth="true"
-OPT adminAuthUser="admin"
-OPT adminAuthPassword="cognos"
-OPT adminAuthPasswordConfirm="cognos"
-OPT runSetupAdmin="true"
-OPT createAdminUserGroup=true
-OPT setupAdminUser="httpusr"
-OPT setupAdminGroup="httpgrp"
-OPT installPlugin="true"
-OPT webserverDefinition="WebServer_LITCIHS1"
-OPT washostname="localhost"
```

3. Start the silent install, as shown in Example 9-11.

Example 9-11   Sample install command for IBM HTTP Server

```
./install -options "responsefile.txt" -silent
```

4. Verify the installation result in the /opt/IBM/HTTPServer/logs/install/log.txt file.

9.3.4 Installing the UpdateInstaller and fix pack for IHS

After the IBM HTTP Server base is installed successfully, continue to install the UpdateInstaller and the fix pack for the IBM HTTP Server:

> **Note:** The installation of the UpdateInstaller can be skipped if you installed it during the WebSphere Application Server installation steps. The same level of UpdateInstaller can be used to install the fix pack for the IBM HTTP Server.

1. Expand the UpdateInstaller package tar file.
2. Go to the UpdateInstaller directory, and modify the responsefile.updiinstaller.txt file, as shown in Example 9-12.

Example 9-12   Sample responsfile.updiindtaller.txt file

```
-OPT silentInstallLicenseAcceptance="true"
-OPT disableOSPrereqChecking="true"
```

```
-OPT installLocation="/opt/IBM/WebSphere/UpdateInstaller"
```

3. Start the UpdateInstaller file, as shown in Example 9-13.

Example 9-13   Sample UpdateInstaller install command with -silent option

```
./install -options "responsefile.updiinstaller.txt" -silent
```

4. After the UpdateInstaller is installed successfully, go to the /opt/IBM/WebSphere/UpdateInstaller/responsefiles directory, and modify the install.txt file to install the fix pack for IBM HTTP Server, as shown in Example 9-14.

> **Note:** Enter the -W option with parameters in one single line.

Example 9-14   Sample responsfiles for installing Fix Pack for IHS

```
-W
maintenance.package="/mnt/WAS7/7.0.0-WS-IHS-LinuxS39064-FP0000009.pak"
-W product.location="/opt/IBM/HTTPServer"
```

5. Run the **update** command to install the fix pack.

> **Note:** Enter the **update** command in one single line.

Example 9-15   Sample update command to install Fix Pack with the -silent option

```
./update.sh -options
"/opt/IBM/WebSphere/UpdateInstaller/responsefiles/install.txt" -silent
```

6. Verify the updateconfig.log file in the /opt/IBM/HTTPServer/logs/update directory.

7. Repeat the install steps for the IBM HTTP Server on other virtual Linux guest.

9.4 Creating server profiles

After WebSphere Application Server installs successfully, we must create the WebSphere Application Server runtime environment with server profiles. To create the necessary server profiles, we used the Silent Install Mode. Two profile types were created for building the Cognos 8 BI runtime infrastructure:

► Deployment Manager Profile
► Managed Profile referred as a Custom Profile or an empty Node

9.4.1 Creating the Deployment Manager profile

Because the server LITWSND1 is the Deployment Manager node of our environment, the following steps are used to create the Deployment Manager profile only on Linux server LITWSND1:

1. Go to the IBM WebSphere Application Server install directory, /opt/IBM/WebSphere/AppServer/bin, and run the **manageprofiles** command to create the Deployment Manager profile, as shown in Example 9-16.

> **Note:** Enter the **manageprofiles** command in one single line.

Example 9-16   Sample Deployment Manager Profile creation command

```
./manageprofiles.sh -create -profileName Dmgr01 -profilePath
"/opt/IBM/WebSphere/AppServer/profiles/Dmgr01" -templatePath
"/opt/IBM/WebSphere/AppServer/profileTemplates/dmgr"
```

2. Verify the profile creation by browsing the logs in the /opt/IBM/WebSphere/AppServer/logs/manageprofiles directory.

9.4.2 Creating the managed profile

To create the managed profile (empty node) in all other Linux servers:

1. Go to the IBM WebSphere Application Server install directory, /opt/IBM/WebSphere/AppServer/bin.

2. Run the **manageprofiles** command to create the managed profile, as shown in Example 9-17.

> **Note:** Enter the **manageprofiles** command in one single line.

Example 9-17   Sample Managed Profile creation command

```
./manageprofiles.sh -create -profileName Appsrv01 -profilePath
"/opt/IBM/WebSphere/AppServer/profiles/Appsrv01" —templatePath
"/opt/IBM/WebSphere/AppServer/profileTemplates/managed"
```

3. Verify the profile creation by browsing the logs in the /opt/IBM/WebSphere/AppServer/logs/manageprofiles directory.

4. After the Managed Profile creates successfully, the profile must be federated to the Deployment Manager node.

5. Go to the /opt/IBM/WebSphere/AppServer/profiles/Appsrv01/bin directory.

6. Run the **addNode** command to federate to the Deployment Manager, as shown in Example 9-18.

Example 9-18   Sample addNode command

```
./addNode.sh 192.168.71.97 8879
```

7. Run the **stopNode**, **syncNode**, and **startNode** command to further verify whether the node was created and federated successfully, as shown in Example 9-19.

Example 9-19   Sample stopNode, syncNode and startNode command

```
./stopNode.sh
./syncNode.sh 192.168.71.97 8879
./startNode.sh
```

9.5 Creating clusters and application servers

After the server profiles are created, application servers or JVM instances, are needed for Cognos. Multiple application clusters and servers are created and used to deploy the Cognos 8 BI components: Content Manager, Report Server, and Gateway.

A cluster is created for each Cognos 8 BI component to support failover and workload balancing capabilities. However, the Cognos Content Manager only uses one active WebSphere Application Server server. The other server remains shutdown for standby purposes.

Due to our lab network setup and configuration, we deployed a copy of the LDAP administration web application on our Linux server to access the Tivoli Directory Server.

We also deployed our onboarding application to demonstrate the onboarding process, which we describe in further detail in Chapter 16, "Onboarding" on page 253.

Table 9-4 on page 149 shows the clusters and application servers being created for the Cognos 8 BI runtime infrastructure.

Table 9-4   Clusters and servers for Cognos 8 BI

Cluster name	Server name	Host
Cognos_content_mgr_cluster	Cognos_Content_Mgr	LITCCM01 LITCCM02
Cognos_gateway_cluster	Cognos_Gateway_server	LITCIHS1 LITCIHS2
Cognos_report_cluster	Cognos_Report_server	LITCRP01 LITCRP02
IDS_adm_client_cluster	IDS_Admin_GUI	LITCBA01 LITCBA02
OnBoarding_tool_cluster	OnBoarding	LITCBA01 LITCBA02

To create the required clusters and application servers:

1. Login to the WebSphere Application Server Admin Console.

2. Create the Clusters and Servers as listed in Table 9-4. Click **Servers** → **Clusters** → **WebSphere application server clusters** → **New**.

3. After servers are created, click **Servers** → **Server Types** → **WebSphere application servers** → **<server_name>** → **Process definition** → **Logging and Tracing** → **IBM Service Logs**.

4. Modify the IBM Service Logs file name as ${SERVER_LOG_ROOT}/activity.log

9.5.1 Modifying the log file for the node agent and Deployment Manager

To have the IBM Services Log separated for each server, modify the location of the IBM Services Log:

1. Login to the WebSphere Application Server Admin Console.

2. Click **System administration** → **Node agents** → **<nodeagent_name>** → **Process definition** → **Logging and Tracing** → **IBM Service Logs**.

3. Modify the node agent log as $(LOG_ROOT)/$(SERVER)/activity.log.

4. Click **Deployment manager** → **Process definition** → **Logging and Tracing** → **IBM Service Logs**.

5. Modify the Deployment Manager log as $(LOG_ROOT)/dmgr/activity.log.

9.6 Creating the IBM HTTP Web Server

The IBM HTTP Server is required for running Cognos 8 BI Gateway components. It is installed on Linux guests LITCIHS1, LITCHIS2, and LITCCM01. The HTTP server on LITCCM01 will be first used for the Cognos installation that has all components on one guest. It is not used at a later stage when the Cognos Gateway server is moved to server LITCIHS1 and LITCIHS2.

> **TIP:** Make sure the HTTP admin server is started. Use the following command to start the admin server if it is not running:
>
> `./opt/IBM/HTTPServer/bin/adminctl start`

To create the required HTTP web servers:

1. Login to the WebSphere Application Server Admin console.

2. Click **System administration** → **Nodes** → **Add Node**.

3. Check the unmanaged node option, and click **Next**.

4. Enter the node name, host name, and select the platform type, as shown in Table 9-5.

Table 9-5   Unmanaged Node for web servers

Node name	Host name	Platform
Unmanaged_node_LITCCM01	LITCCM01	Linux
Unmanaged_node_LITCIHS1	LITCIHS1	Linux
Unmanaged_node_LITCIHS2	LITCIHS2	Linux

5. To add new web servers, click **Servers** → **Server Types** → **Web servers** → **New**.

6. Use the information in Table 9-6 to create the web server for LITCIHS1.

Table 9-6   Web server details on LITCIHS1

Field	Value
Node	Unmanaged_node_LITCIHS1
Server Name	WebServer_LITCIHS1
Type	IBM IHS Server
Template	IHS

Field	Value
Port	80
Installation Location	/opt/IBM/HTTPServer
Configuration file name	${WEB_INSTALL_ROOT}/conf/httpd.conf
Plug-in installation location	/opt/IBMHttpServer/Plugins
Application mapping to the Web server	All
Administration Server port	8008
Username	admin
Password	cognos

7. Create the following two directories, if they do not exist. The directory is used for the WebSphere Application Server plug-in set up and configuration steps:

```
/opt/IBM/HTTPServer/Plugins/logs/WebServer_LITCIHS1
/opt/IBM/HTTPServer/Plugins/config/WebServer_LITCIHS1
```

8. Verify the directive, WebSpherePluginConfig, in the /opt/IBM/HTTPServer/conf/httpd.conf file, as follows:

Note: Enter the parameter in one single line.

```
WebSpherePluginConfig
/opt/IBM/HTTPServer/Plugins/config/WebServer_LITCIHS1/plugin-cfg.xml
```

9. Repeat the steps to create the web servers on LITCIHS2 and LITCCM01.

9.7 Enabling security

Security is important in a production environment. By enabling security, we protect the server from unauthorized users and provide application isolation and requirements for authenticating application users.

Section 9.7.1, "Setting up global security" on page 152 shows the steps to enable Global Security for WebSphere Application Server.

Section 9.7.2, "Setting up SSL" on page 153 shows the procedures for setting up SSL for secure connection.

9.7.1 Setting up global security

In our lab environment, we used an existing stand-alone Lightweight Directory Access Protocol (LDAP) server for a user registry/repository. It authenticates users and retrieves user and group information to perform security-related functions, including authentication and authorization.

> **TIP:** Always make a backup copy of the security.xml file, which is in the /opt/IBM/WebSphere/AppServer/profiles/Dmgr01/config/cells/LITWSND1Cell 01 directory of the Deployment Manager node.

To enable Global Security:

1. Login to the WebSphere Application Server Admin console.
2. Click **Security** → **Global security** → **Security Configuration Wizard**.
3. Use the information from Table 9-7 to set up global security.

> **Note:** The desired primary administrative user ID must be existing or created in the repository ahead of time.

Table 9-7   Values for Security Configuration Wizard

Field	Value
Enable application security	Select this option
Use Java 2 security	Do not select this option
Select user repository	Stand-alone LDAP registry
Primary administrative user name	wasadmin2
Type of LDAP Server	IBM Tivoli Directory Server
Host	192.168.71.114
Port	389
Base distinguished name (DN)	ou=People,o=ticl,c=us
Bind distinguished name (DN)	cn=root
Bind password	linux390

Figure 9-2 on page 153 shows the values to enable security.
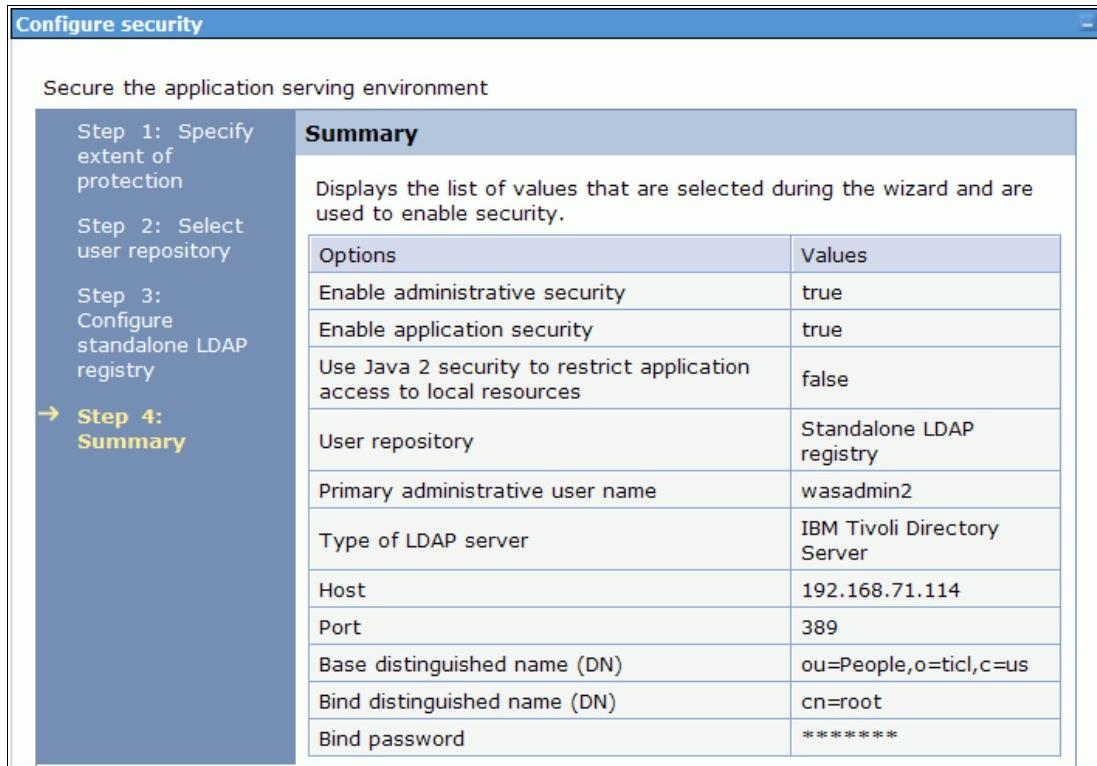
Figure 9-2   Configure Security

9.7.2 Setting up SSL

Instead of using the shipped default key database and certificates, we created new key databases with our own self-signed certificates following best practices for a production environment. SSL is enabled and uses the newly created key databases and certificates. In this section, we describe the steps that enable the SSL connection.

Creating SSL key databases

We created five SSL key databases using the ikeyman utility. Table 9-8 on page 154 shows the file names and the corresponding file type used in our lab environment.

TIP: IBM WebSphere Application Server is shipped with a version of the ikeyman utility in the *<install_root>*/bin directory.

Table 9-8   SSL key database

File name	File type	Password
itsoserverkey.p12	PKCS12	key4was
itsoservertrust.p12	PKCS12	key4was
itsoclientkey.p12	PKCS12	key4was
itsoclienttrust.p12	PKCS12	key4was
itso_plugin.kdb	CMS	key4was
itso_plugin.sth	Stash file	

To create the necessary SSL key databases:

1. Create itsoserverkey.p12 and a self-signed certificate, and export the certificate as an arm file.

2. Create itsoclientkey.p12 and a self-signed certificate, and export the certificate as an arm file.

3. Create itso_plugin.kdb and a self-signed certificate, generate the stash file, and export the certificate as an arm file.

4. Create itsoservertrust.p12, and import the three arm files created in the previous steps.

5. Create itsoclienttrust.p12. Import the three arm files that you exported in the previous steps.

6. Create a new directory /opt/IBM/WebSphere/sslkey in all Linux guests and copy all SSL key databases previously created to this directory.

Example 9-20 shows the SSL key database files on Linux guest LITWSND1.

Example 9-20   SSL key database files on Linux guest LITWSND1

```
LITWSND1:/opt/IBM/WebSphere/sslkey # ll
total 256
-rw-r--r-- 1 root root     80 Apr 12 11:32 itso_plugin.crl
-rw-r--r-- 1 root root 130080 Apr 12 11:32 itso_plugin.kdb
-rw-r--r-- 1 root root     80 Apr 12 11:32 itso_plugin.rdb
-rw-r--r-- 1 root root    129 Apr 12 11:32 itso_plugin.sth
-rw-r--r-- 1 root root  27162 Apr 12 11:32 itsoclientkey.p12
-rw-r--r-- 1 root root  27842 Apr 12 11:32 itsoclienttrust.p12
-rw-r--r-- 1 root root  28602 Apr 12 11:32 itsoserverkey.p12
-rw-r--r-- 1 root root  27842 Apr 12 11:32 itsoservertrust.p12
```

Creating the SSL Key store

To create the SSL Key store:

1. Login to the WebSphere Application Server Admin console.

2. Click **Security** $\rightarrow$ **SSL certificate and key management** $\rightarrow$ **Key stores and certificates** $\rightarrow$ **New**.

3. Use the values in Table 9-9 to create the Server Key Store.

Table 9-9   Server Key store

Field	Value
Name	ITSO_Server_Key_Store
Path	/opt/IBM/WebSphere/sslkey/itsoserverkey.p12
Password	key4was
Type	PKCS12

4. Use the values in Table 9-10 to create the Trust Key Store.

Table 9-10   Trust Key Store

Field	Value
Name	ITSO_Server_Trust_Store
Path	/opt/IBM/WebSphere/sslkey/itsoservertrust.p12
Password	key4was
Type	PKCS12

Creating SSL settings

To create SSL settings:

1. Login to the WebSphere Application Server Admin console.

2. Click **Security** $\rightarrow$ **SSL certificate and key management** $\rightarrow$ **SSL Configurations** $\rightarrow$ **New**.

3. In the Name field, type `ITSO_WAS_SSL_Settings`.

4. Select the newly created Key Store and Trust Store.

5. Click **Get certificate aliases**.

6. In the Default server certificate alias and the Default client certificate alias fields, select **websphere server key**.

7. Click **OK** to save the settings.

Figure 9-3 shows the SSL settings that we used.
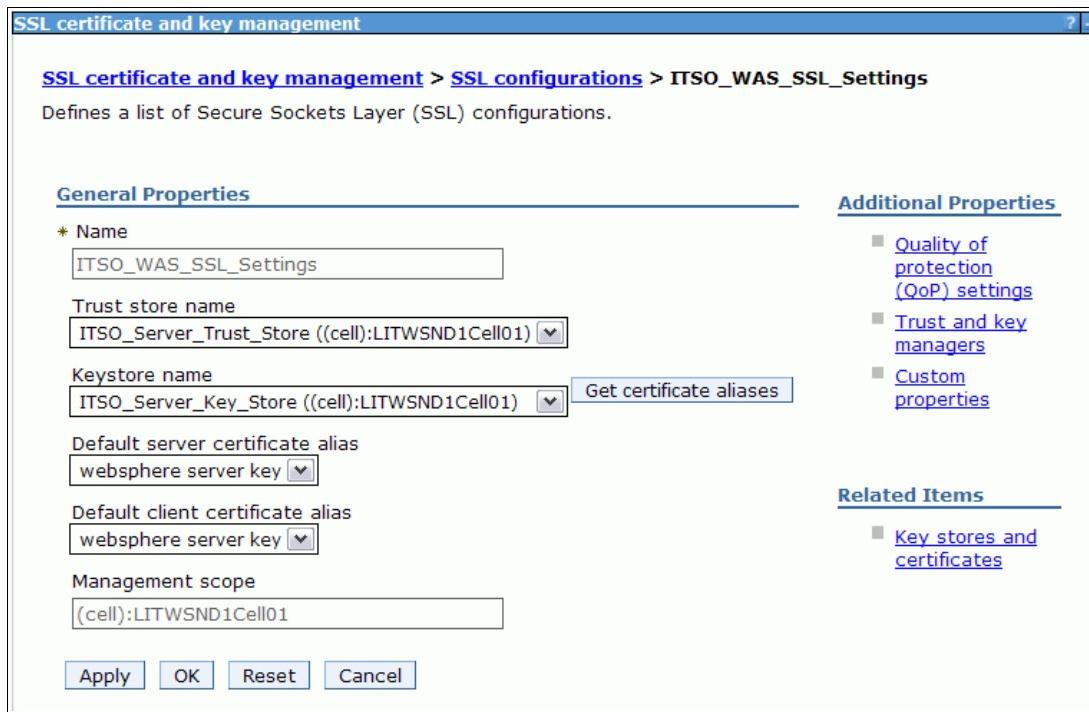
Figure 9-3   SSL settings

Modifying the endpoint security configuration at the cell level

To modify the endpoint security configuration at the cell level for both Inbound and Outbound:

1. Login to the WebSphere Application Server Admin console.

2. Click **Security** → **SSL certificate and key management**.

3. Under Configuration settings, near the bottom of the window, click → **Manage endpoint security configurations**.

4. Click the Cell Link under Inbound.

5. In the SSL configuration field, select **ITSO_WAS_SSL_Settings**.

6. In the Certificate alias in key store field, select **websphere server key**.

7. Repeat the steps for the Outbound endpoint at the Cell Level.

Figure 9-4 on page 157 shows the endpoint security configuration for Inbound at the cell level.
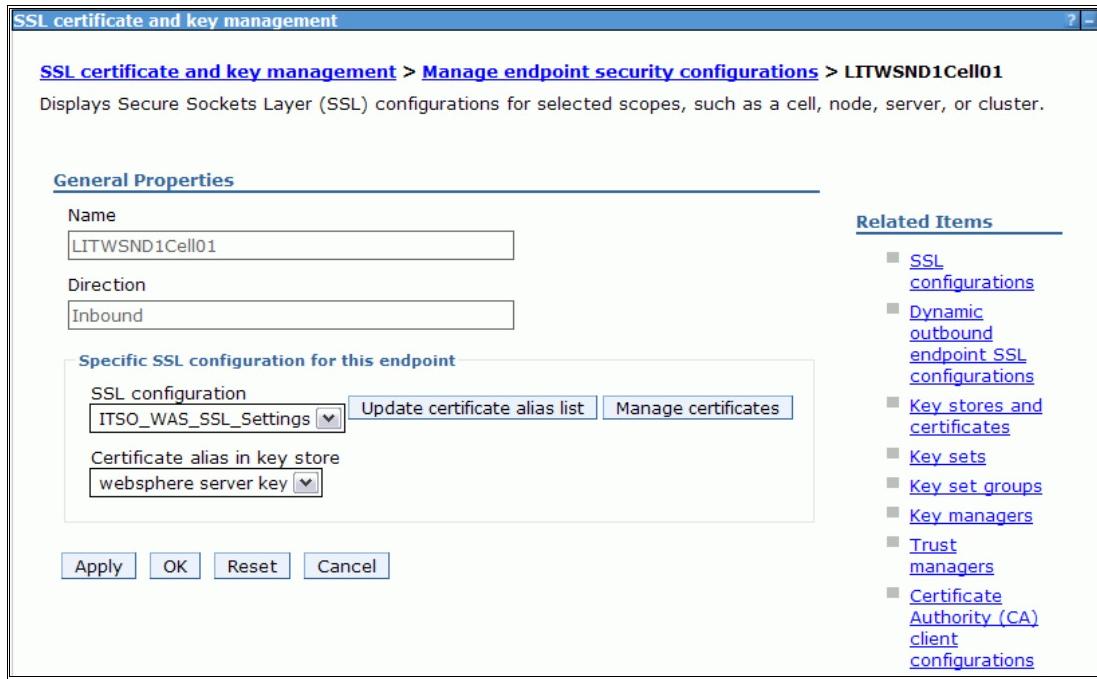
Figure 9-4   End point security configuration for Inbound at the cell level

Modifying the endpoint security configuration at the node level

To modify the endpoint security configuration at the node level for both Inbound and Outbound:

1. Login to the WebSphere Application Server Admin console.

2. Click **Security** → **SSL certificate and key management**.

3. Under Configuration settings, near bottom of the panel, click → **Manage endpoint security configurations**.

4. Click the first node link under Inbound.

5. Deselect the **Override inherited values** option.

6. Repeat these steps for all managed node links under Inbound.

7. Repeat these steps for all managed node links under Outbound.

Figure 9-5 on page 158 shows the endpoint security configuration for Inbound at the node level.
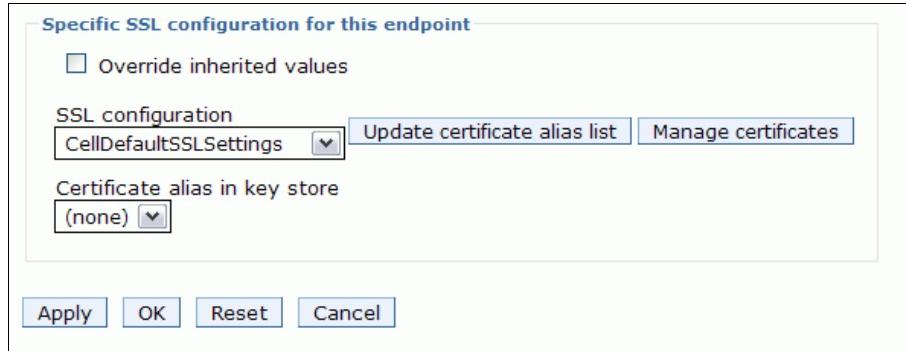
Figure 9-5   End point security configuration for Inbound at the Node Level

Modifying the Deployment Manager configuration

To modify the Deployment Manager set up:

1. Login to the WebSphere Application Server Admin console.

2. Click **System Administration** → **Deployment Manager** → **Web container transport chains** → **WCInboundAdminSecure**.

3. Click **SSL Inbound Channel (SSL 1)**.

4. In the SSL Configuration section, select Specific to this endpoint and ITSO_WAS_SSL_Settings.

Figure 9-6 on page 159 shows the values that we selected for the Deployment Manager Transport Chain.

Figure 9-6   Deployment Manager WCInboundAdminSecure settings

Preparing symbolic links for the plug-in database

To prepare symbolic links for the plug-in database:

1. Login to the Deployment Manager Linux guest.

2. Go to the
   /opt/IBM/WebSphere/AppServer/profiles/Dmgr01/config/cells/<cell_name>/n
   odes/unmanaged_node01_<hostname>/servers/webserver_<hostname>/
   directory.

3. Create the following two symbolic links:

   ```
   ln -s /opt/IBM/WebSphere/sslkey/itso_plugin.kdb itso_plugin.kdb
   ln -s /opt/IBM/WebSphere/sslkey/itso_plugin.sth itso_plugin.sth
   ```

4. Repeat these steps for all configured web servers.

Updating the web server configuration

To update the web server configuration:

1. Login to the Admin console, and click **Servers** → **Server Types** → **Web servers**.

2. Click the desired Web Server link.

3. Enter `443` as the Port number.

4. Save the changes, and repeat the update to all other web servers.

5. Login to LITCIHS1, LITCIHS2, and LITCCM01.

6. Go to the /opt/IBM/HTTPServer/conf directory.

7. Modify the httpd.conf file, and add the lines shown in Example 9-21.

Example 9-21   httpd.conf changes

```
LoadModule ibm_ssl_module modules/mod_ibm_ssl.so
Listen 443
Keyfile /opt/IBM/WebSphere/sslkey/itso_plugin.kdb
SSLEnable
SSLServerCert WebSphere Plugin key
```

Updating the web server plug-in configuration

To update the web server plug-in configuration:

1. Login to the Admin console, and click **Servers** → **Server Types** → **Web servers**.

2. Click the desired web server link.

3. Under Additional Properties, click **Plug-in properties**.

4. Update the plug-in key store file name to `itso_plugin.kdb`.

5. Save the changes and come back to this page again.

6. Click **Copy to Web server key store directory**.

7. Generate plug-in and propagate plug-in for the updated web server.

8. Restart the web server.

9. Repeat these steps for all configured web servers.

Figure 9-7 on page 161 shows the web server plug-in configuration that we used.

Figure 9-7   Web server plug-in configuration

Modifying the Dmgr soap.client.props file

To modify the Dmgr soap.client.props file:

1. Login to the Deployment Manager Linux guest.

2. Go to the /opt/IBM/WebSphere/AppServer/profiles/Dmgr01/properties directory.

3. Modify the soap.client.props file using the values specified in Table 9-11 on page 162.

4. Go to the directory /opt/IBM/WebSphere/AppServer/profiles/Dmgr01/bin.

5. Run ./PropFilePasswordEncoder.sh /opt/IBM/WebSphere/AppServer/profiles/Dmgr01/properties/soap.client.props com.ibm.SOAP.loginPassword.

6. Restart the Deployment Manager.

Table 9-11   values changed in soap.client.props file

Parameter	Values
com.ibm.SOAP.securityEnabled	true
com.ibm.SOAP.loginUserid	wasadmin2
com.ibm.SOAP.loginPassword	cogn0s00

Modifying the Dmgr ssl.client.props file

To modify the Dmgr ssl.client.props file:

1. Login to the Deployment Manager Linux guest.

2. Go to the directory /opt/IBM/WebSphere/AppServer/profiles/Dmgr01/properties.

3. Modify file ssl.client.props using the values specified in Table 9-12.

4. Go to the directory /opt/IBM/WebSphere/AppServer/profiles/Dmgr01/bin.

5. Run ./PropFilePasswordEncoder.sh /opt/IBM/WebSphere/AppServer/profiles/Dmgr01/properties/ssl.client.props com.ibm.ssl.keyStorePassword.

6. Run ./PropFilePasswordEncoder.sh /opt/IBM/WebSphere/AppServer/profiles/Dmgr01/properties/ssl.client.props com.ibm.ssl.trustStorePassword.

7. Restart the Deployment Manager.

Table 9-12   Values changed in soap.client.props file

Parameter	Values
com.ibm.ssl.keyStore	/opt/IBM/WebSphere/sslkey/itsoclientkey.p12
com.ibm.ssl.keyStorePassword	key4was
com.ibm.ssl.trustStore	/opt/IBM/WebSphere/sslkey/itsoclienttrust.p12
com.ibm.ssl.trustStorePassword	key4was

Modifying the AppServer soap.client.props file

To modify the AppServer soap.client.props file:

1. Login to each WebSphere Application Server application Linux guest.

2. Go to the directory
   /opt/IBM/WebSphere/AppServer/profiles/ApsSrv01/properties.

3. Modify the soap.client.props file using the values specified in Table 9-13.

4. Go to the directory /opt/IBM/WebSphere/AppServer/profiles/Appsrv01/bin.

5. Run ./PropFilePasswordEncoder.sh
   /opt/IBM/WebSphere/AppServer/profiles/Appsrv01/properties/soap.client.pro
   ps com.ibm.SOAP.loginPassword.

6. Restart all node agent and application servers.

Table 9-13   Values changed in soap.client.props file

Parameter	Values
com.ibm.SOAP.securityEnabled	true
com.ibm.SOAP.loginUserid	wasadmin2
com.ibm.SOAP.loginPassword	cogn0s00

Modifying the AppServer ssl.client.props file

To modify the AppServer ssl.client.props file:

1. Login to each WebSphere Application Server application Linux guest.

2. Go to the directory
   /opt/IBM/WebSphere/AppServer/profiles/Appsrv01/properties.

3. Modify file ssl.client.props using the values specified in Table 9-14 on
   page 164.

4. Go to the directory /opt/IBM/WebSphere/AppServer/profiles/Appsrv01/bin.

5. Run ./PropFilePasswordEncoder.sh
   /opt/IBM/WebSphere/AppServer/profiles/Appsrv01/properties/ssl.client.props
   com.ibm.ssl.keyStorePassword.

6. Run ./PropFilePasswordEncoder.sh
   /opt/IBM/WebSphere/AppServer/profiles/Appsrv01/properties/ssl.client.props
   com.ibm.ssl.trustStorePassword.

7. Restart all node agents and application servers.

Table 9-14   values changed in soap.client.props file

Parameter	Values
com.ibm.ssl.keyStore	/opt/IBM/WebSphere/sslkey/itsoclientkey.p12
com.ibm.ssl.keyStorePassword	key4was
com.ibm.ssl.trustStore	/opt/IBM/WebSphere/sslkey/itsoclienttrust.p12
com.ibm.ssl.trustStorePassword	key4was

9.8 Administration and operations

In this section, we describe some basic WebSphere Application Server administration and operations. We focus on using the command line utility and the *wsadmin tool*. The operations and administration of the IBM WebSphere Application Server are done mainly using the administrative console. The administrative console supports a full range of product administrative activities.

The IBM WebSphere Application Server wsadmin tool supports a full range of product administrative activities, which includes providing the ability to run scripts. Scripting is a non-graphical alternative that can be used to configure and manage the IBM WebSphere Application Server.

For non-interactive or unattended operations and administration activities, such as resources set up and provisioning in the Cloud, post-installation and customization tasks for administrators, deploying applications onto application servers, and administering applications and their server environments, the wsadmin tool and commands are used.

The IBM WebSphere Application Server is shipped with a ready-to-use command line utility for several basic server operations.

9.8.1 Administrative console

The web-based administrative console tool is used to manage the IBM WebSphere Application Server and allows the administrator to perform a full range of product administrative activities. In our lab environment, the administrative console application starts automatically when the Deployment Manager server is started to which the administrative console belongs. To access our Administration Console, we use the URL, as shown in Figure 9-22 on page 165.

Example 9-22   Administrative Console URL

```
http://192.168.71.97:9060/admin
```

9.8.2 Command line utility

There are several IBM WebSphere Application Server command line utilities that you can use to start, stop, and monitor application server processes and nodes. These tools only work on local servers and nodes. They cannot operate on a remote server or node. All IBM WebSphere Application Server supplied command line utilities can be found in the <*WAS_profiles_root*>/bin directory. Table 9-15 shows several commonly used commands.

Table 9-15   IBM WebSphere Application Server commonly used command line utility

Utility	Function
addNode.sh	Add a node
backupConfig.sh	Back up configuration
dumpNameSpace.sh	Dump name space
ikeyman.sh	Key database management
manageprofiles.sh	Profile management
removeNode.sh	Remove a node
restoreConfig.sh	Restore configuration from back up
serverStatus.sh	Inquire server status
startManager	Start Deployment Manager
startNode.sh	Start a node
startServer.sh	Start a server
stopManager	Stop Deployment Manager
stopNode.sh	Stop a node
stopServer.sh	Stop a server
syncNode.sh	Full synchronization
versionInfo.sh	Show IBM WebSphere Application Server version installed
wsadmin.sh	Invoke Admin Tool

In Example 9-23 we show some typical commands using the command line utility.

Example 9-23   Command line utility samples

```
./stopServer Cognos_Report_server
./startServer Cognos_Report_server
./serverStatus -all
```

There is a variety of IBM HTTP Server command line utilities that you can use to start, stop, and restart server processes. These tools only work on local servers. They cannot operate on a remote server. All IBM HTTP Server-supplied command line utilities are found in the directory <*IHS_install_root*>/bin. Table 9-16 shows several commonly used utilities.

Table 9-16   HTTP server supplied utility

Utility	Function
adminctl	Start / Stop / Restart HTTP Admin server
apachectl	Start / Stop / Restart HTTP server
htpasswd	Generate Admin server password
ikeyman	Key database management
versionInfo.sh	Show IBM HTTP Server version installed

In Example 9-24 we show some typical uses of the HTTP server command line utility.

Example 9-24   Samples HTTP server command line utility

```
./adminctl stop
./adminctl start
./apachectl restart
```

9.8.3 The wsadmin tool with scripting commands

The IBM WebSphere administrative (wsadmin) scripting program is a powerful, non-graphical command interpreter environment that allows you to run administrative operations using a scripting language. Scripting is a non-graphical alternative that can be used to configure and administer the applications and application serving environment. The IBM WebSphere Application Server wsadmin tool provides the ability to run scripts.

The wsadmin tool is intended for unattended operations. It can be used to perform the same tasks that can be performed using the administrative console.

The wsadmin tool supports a full range of product administrative activities. There are five objects that are available for scripting, as shown in Table 9-17. The wsadmin tool only supports Jacl and Jython scripting languages. Jacl is the language specified by default. If Jython scripting language is desire, use the `-lang` option or specify it in the wsadmin.properties file. The wsadmin scripting client can be started interactively, as an individual command, in a script, or in a profile.

Table 9-17   wsadmin management objects

Object	Function
AdminControl	Use to run operational commands
AdminConfig	Use to run configurational commands to create or modify WebSphere Application Server configurational elements
AdminApp	Use to administer applications
AdminTask	Use to run administrative commands
Help	Use to obtain general help

Example 9-25 shows some sample help commands.

Example 9-25   Sample help commands

```
print AdminLibHelp.help()
print AdminApp.help()
print AdminTask.help("-commands","create*")
```

Sample scripts used for our lab setup

Example 9-26 shows how we used the wsadmin script to create our testing WebSphere Application Server environment for Cognos 8 BI.

Example 9-26   Invoke wsadmin tool

```
wsadmin -lang jython -f /itso/create_servers.py
wsadmin -lang jython -f /itso/config_ssl.py
```

Figure 9-27 on page 168 shows the first input file content for the wsadmin script.

Note: All Management object commands must be entered in one line.

Example 9-27   Sample create_servers.py file

```
AdminServerManagement.createApplicationServer("LITCCM01Node01",
"Cognos_Content_Mgr", "default")
AdminClusterManagement.createClusterWithFirstMember("Cognos_content_Mgr
_cluster", "APPLICATION_SERVER", "LITCCM01Node01",
"Cognos_Content_Mgr")
AdminClusterManagement.createClusterMember("Cognos_content_Mgr_cluster"
, "LITCCM02Node01", "Cognos_Content_Mgr")

AdminServerManagement.createApplicationServer("LITCRP01Node01",
"Cognos_Report_server", "default")
AdminClusterManagement.createClusterWithFirstMember("Cognos_report_clus
ter", "APPLICATION_SERVER", "LITCRP01Node01", "Cognos_Report_server")
AdminClusterManagement.createClusterMember("Cognos_report_cluster",
"LITCRP02Node01", "Cognos_Report_server")

AdminServerManagement.createApplicationServer("LITCIHS1Node01",
"Cognos_Gateway_server", "default")
AdminClusterManagement.createClusterWithFirstMember("Cognos_gateway_clu
ster", "APPLICATION_SERVER", "LITCIHS1Node01", "Cognos_Gateway_server")
AdminClusterManagement.createClusterMember("Cognos_gateway_cluster",
"LITCIHS2Node01", "Cognos_Gateway_server")

AdminServerManagement.createApplicationServer("LITCBA01Node01",
"IDS_Admin_GUI", "default")
AdminClusterManagement.createClusterWithFirstMember("IDS_admin_client_c
luster", "APPLICATION_SERVER", "LITCBA01Node01", "IDS_Admin_GUI")
AdminClusterManagement.createClusterMember("IDS_admin_client_cluster",
"LITCBA02Node01", "IDS_Admin_GUI")

AdminServerManagement.createApplicationServer("LITCBA01Node01",
"OnBoarding", "default")
AdminClusterManagement.createClusterWithFirstMember("OnBoarding_tool_cl
uster", "APPLICATION_SERVER", "LITCBA01Node01", "OnBoarding")
AdminClusterManagement.createClusterMember("OnBoarding_tool_cluster",
"LITCBA02Node01", "OnBoarding")

AdminTask.createUnmanagedNode('[-nodeName Unmanaged_node_LITCCM01
-hostName LITCCM01 -nodeOperatingSystem linux]')

AdminConfig.save()

AdminTask.createWebServer('Unmanaged_node_LITCCM01', '[-name
WebServer_LITCCM01 -templateName IHS -serverConfig [-webPort 80
```

```
-serviceName -webInstallRoot /opt/IBM/HTTPServer -webProtocol HTTP
-configurationFile -errorLogfile -accessLogfile -pluginInstallRoot
/opt/IBM/HTTPServer/Plugins -webAppMapping ALL] -remoteServerConfig
[-adminPort 8008 -adminUserID admin -adminPasswd cognos -adminProtocol
HTTP]]')

AdminConfig.save()

AdminTask.createUnmanagedNode('[-nodeName Unmanaged_node_LITCIHS1
-hostName LITCIHS1 -nodeOperatingSystem linux]')

AdminConfig.save()

AdminTask.createWebServer('Unmanaged_node_LITCIHS1', '[-name
WebServer_LITCIHS1 -templateName IHS -serverConfig [-webPort 80
-serviceName -webInstallRoot /opt/IBM/HTTPServer -webProtocol HTTP
-configurationFile -errorLogfile -accessLogfile -pluginInstallRoot
/opt/IBM/HTTPServer/Plugins -webAppMapping ALL] -remoteServerConfig
[-adminPort 8008 -adminUserID admin -adminPasswd cognos -adminProtocol
HTTP]]')

AdminConfig.save()

AdminTask.createUnmanagedNode('[-nodeName Unmanaged_node_LITCIHS2
-hostName LITCIHS2 -nodeOperatingSystem linux]')

AdminConfig.save()

AdminTask.createWebServer('Unmanaged_node_LITCIHS2', '[-name
WebServer_LITCIHS2 -templateName IHS -serverConfig [-webPort 80
-serviceName -webInstallRoot /opt/IBM/HTTPServer -webProtocol HTTP
-configurationFile -errorLogfile -accessLogfile -pluginInstallRoot
/opt/IBM/HTTPServer/Plugins -webAppMapping ALL] -remoteServerConfig
[-adminPort 8008 -adminUserID admin -adminPasswd cognos -adminProtocol
HTTP]]')

AdminConfig.save()

AdminServerManagement.startAllServers("LITCCM01Node01")
AdminServerManagement.startAllServers("LITCCM02Node01")
AdminServerManagement.startAllServers("LITCRP01Node01")
AdminServerManagement.startAllServers("LITCRP02Node01")
AdminServerManagement.startAllServers("LITCBA01Node01")
AdminServerManagement.startAllServers("LITCBA02Node01")
AdminServerManagement.startAllServers("LITCIHS1Node01")
```

```
AdminServerManagement.startAllServers("LITCIHS2Node01")
```

Example 9-28 shows the 2nd input file content for the wsadmin script.

> **Note:** All Management object commands must be entered in one line.

Example 9-28   Sample config_ssl.py file

```
AdminTask.createKeyStore('[-keyStoreName ITSO_Server_Key_Store
-keyStoreType PKCS12 -keyStoreLocation
/opt/IBM/WebSphere/sslkey/itsoserverkey.p12 -keyStorePassword key4was
-keyStorePasswordVerify key4was -keyStoreDescription SSL Server Key
Store for ITSO]')

AdminTask.createKeyStore('[-keyStoreName ITSO_Server_Trust_Store
-keyStoreType PKCS12 -keyStoreLocation
/opt/IBM/WebSphere/sslkey/itsoserverkey.p12 -keyStorePassword key4was
-keyStorePasswordVerify key4was -keyStoreDescription SSL Server Key
Store for ITSO]')

AdminConfig.save()

AdminTask.createSSLConfig('[-alias ITSO_WAS_SSL_Settings -type JSSE
-scopeName (cell):LITWSND1Cell01 -keyStoreName ITSO_Server_Key_Store
-keyStoreScopeName (cell):LITWSND1Cell01 -trustStoreName
ITSO_Server_Trust_Store -trustStoreScopeName (cell):LITWSND1Cell01
-serverKeyAlias "websphere server key" -clientKeyAlias "websphere
server key" ]')

AdminConfig.save()

AdminTask.modifySSLConfigGroup('[-name LITWSND1Cell01 -direction
inbound -certificateAlias "websphere server key" -scopeName
(cell):LITWSND1Cell01 -sslConfigAliasName ITSO_WAS_SSL_Settings
-sslConfigScopeName (cell):LITWSND1Cell01 ]')

AdminTask.modifySSLConfigGroup('[-name LITWSND1Cell01 -direction
outbound -certificateAlias "websphere server key" -scopeName
(cell):LITWSND1Cell01 -sslConfigAliasName ITSO_WAS_SSL_Settings
-sslConfigScopeName (cell):LITWSND1Cell01 ]')

AdminConfig.save()
```

More information about the wsadmin script command and library are at:

http://publib.boulder.ibm.com/infocenter/wasinfo/v7r0/index.jsp?topic=/
com.ibm.websphere.base.doc/info/aes/ae/txml_7scrlibhelp.html

9.8.4 The wsadmin tool with properties files

Property files can be used to manage the environment and configuration objects. Some of the configuration objects cannot be modified through commands or the script library, but can only be modified using properties files. The process is simple. First extract the configuration objects to simple properties file format, modify the extracted properties file, and apply the modified properties file to update the system configuration. To modify configuration objects using properties files:

1. Use AdminTask.extractConfigProperties to extract the object properties to a flat file.
2. Edit the extracted properties file with the desired values.
3. Use AdminTask.applyConfigProperties to update the object properties.
4. Save the updates using AdminConfig.save().

Example 9-29 shows updating the node agent service log location.

> **Note:** The AdminTask command must be entered in one line.

Example 9-29   Update the node agent service log directory location

```
./wsadmin.sh -lang jython
AdminTask.extractConfigProperties('[-propertiesFileName
/itso/itso-LITCIHS1Node01-servicelog.props -configData
Node=LITCIHS1Node01 -selectedSubTypes [ServiceLog]]')

...exit from wsadmin, edit the output props file and invoke wsadmin
again to continue..

./wsadmin.sh -lang jython
AdminTask.applyConfigProperties('[-propertiesFileName
/itso/itso-LITCIHS1Node01-servicelog.props]')

AdminConfig.save()
```

10

Cognos installation

This chapter describes the installation process for IBM Cognos BI 8.4.1.

IBM Cognos BI 8.4.1 consists of server and client components. The server components are the Content Manager, Content Manager Database, Application Tier components, and the Gateway Server. In our lab environment, the server components are installed on Linux guests.

The client components that we use in the lab environment are IBM Cognos Connection as the portal for Query Studio, Analysis Studio, Event Studio, and Report Studio. The client components are accessed with a web browser.

The metamodel for the reports is created in Framework Manager, which runs on a Windows 2003 installation in our lab environment.

To provide a scalable and robust environment for a Smart Analytics Cloud, the server components are installed on separate guests in a distributed configuration. Figure 10-1 on page 174 gives an overview of the components in the lab environment. Figure 10-2 on page 175 shows the connection between the DB2 server and the DB2 clients of each Cognos component.

173
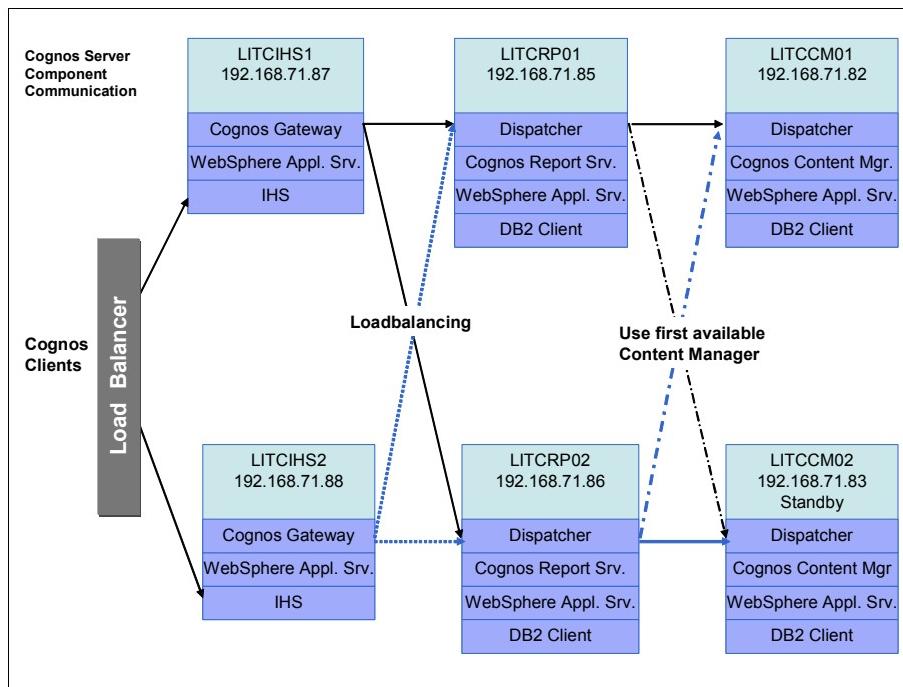
Figure 10-1   Server overview of installed Cognos components

The following sections detail the installation and configuration options that we selected for our cloud environment. Complete installation instructions are available in the *IBM Cognos 8 Business Intelligence Installation and Configuration Guide for Linux on* System z.

The first section gives a brief description of the required components (they are already described in Chapter 5, "Architecture overview" on page 47). Section 10.1, "Required components and prerequisites" on page 175 lists the prerequisites of the installation. The installation itself is explained in the remainder of this chapter, beginning with 10.2, "Installing and configuring Cognos components" on page 176.

By default, Cognos installs and uses Tomcat as the application server for the Java components. In 10.5, "Configuring the IBM Cognos 8 to run with IBM WebSphere Application Server" on page 190, we describe why we chose to use IBM WebSphere Application Server instead of Tomcat and how that configuration WebSphere Application Server done.

Figure 10-2 on page 175 shows the Cognos Database communication.
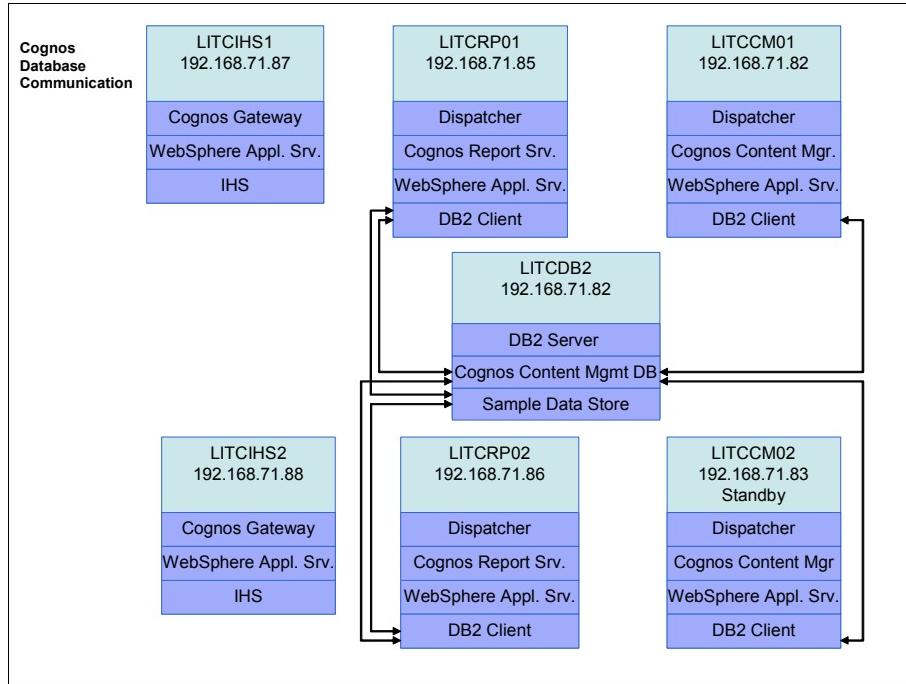
Figure 10-2   Cognos Database communication

10.1 Required components and prerequisites

The basic components in the IBM Cognos BI 8.4.1 server installation are:

- ► Content Manager
- ► Report Server
- ► Gateway Server

Figure 10-1 on page 174 shows the target scenario for the IBM Cognos BI 8.4.1 installation. Each server component is installed on a separate guest. Each component has a clustered or standby component for high availability. The load balancer that is drawn left to the Cognos Gateway is not installed in our lab environment, but must be implemented in a real world high availability solution to distribute requests to the Web servers.

Before the IBM Cognos BI 8.4.1 installation can be started prerequisites must be completed:

1. Install DB2 on the database server LITCDB2.

2. Create the Content Manager database. In our distributed environment, we installed a Content Manager database named contstor on the DB2 database server LITCDB2 by running the script C8DB2.sh, which is supplied with the Cognos installation code.

3. Install DB2 9.5 runtime clients on all Linux guests that will have IBM Cognos BI 8.4.1 Content Manager or Application Tier components. The client is required to provide connectivity to the content store and query databases.

4. Install IBM WebSphere Application Server on all Linux guests that will have IBM Cognos BI 8.4.1 server components. By default, Cognos installs with a Tomcat application server. After installing Cognos, we migrated it from Tomcat to WebSphere Application server, as described in 10.5, "Configuring the IBM Cognos 8 to run with IBM WebSphere Application Server" on page 190.

5. Install the web server software. In our lab environment, we installed IBM HTTP Server on LITCIHS1 and LITCIHS2.

6. Ensure X Windows functionality is available on the Content Manager, Application Tier, and Gateway servers because it is required to run the Cognos Configuration GUI.

7. Set the JAVA_HOME environment variable for the ID that is used to install and configure Cognos. We used root in our environment and set JAVA_HOME=/usr/lib64/jvm/jre.

10.2 Installing and configuring Cognos components

The installation in the sample is performed as root. However, a dedicated user with the appropriate permissions can also be used for installation. Cognos components must be installed in the following order: Content Manager, Application Tier, and finally Gateway.

10.2.1 Installing and configuring Content Manager

After completing the prerequisites, as described in 10.1, "Required components and prerequisites" on page 175, we installed the primary Content Manager on LITCCM01 and the standby on LITCCM02. To install and configure the Content Manager:

1. Open a terminal session through X Windows on the server where the Content Manager is to be installed. Go to the directory where the Cognos installation code resides, and start the Cognos setup program with ./issetup.

2. Follow the steps of the installation wizard. In the Component selection window, Figure 10-3, select only the Content Manager component, and click **Next**.
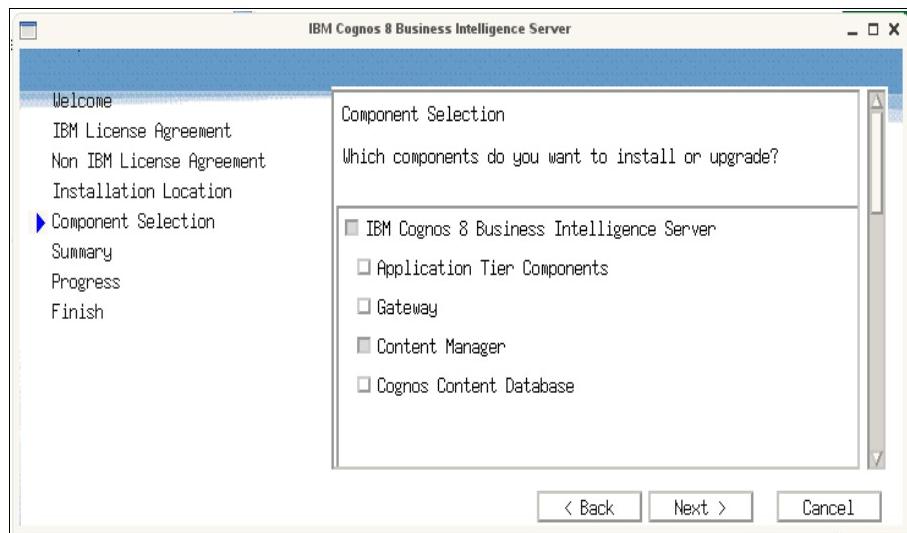
Figure 10-3   Component selection during Cognos installation

3. Verify the component and directory selections, and then follow the prompts to complete the installation.

4. After installation, update the LD_LIBRARY_PATH environment variable to include the appropriate library paths for the Cognos and DB2 versions installed. In our installation, we set this value in the .profile of the ID that is running Cognos, as shown in Example 10-1.

Example 10-1   Sample LD_LIBRARY_PATH entry in .profile

```
export
LD_LIBRARY_PATH=/home/cogdba/sqllib/lib64:/opt/cognos/c8_64/bin64:/opt
/cognos/c8_64/cgi-bin
```

5. Copy the DB2 JDBC drivers from the DB2 client installation to the Cognos installation. In our installation, we copied db2jcc.jar, db2jcc4.jar, db2java.jar and db2jcc_license_cu.jar files from /opt/ibm/db2/V9.5/java to /opt/cognos/c8_64/webapps/p2pd/WEB-INF/lib.

6. Ensure that the JAVA_HOME environment variables are set to the Java Runtime Environment (JRE) location for your installation, as shown in Example 10-2 on page 178.

Example 10-2   Sample JAVA_HOME environment variable value

```
export JAVA_HOME=/usr/lib64/jvm/jre-1.6.0-ibm
```

7. Copy the Cognos 8 cryptography file to the JRE. We copied bcprov-jdk14-134.jar from /opt/cognos/c8_64/bin64/jre/1.5.0/lib/ext to /usr/lib64/jvm/jre-1.6.0-ibm/lib/ext/.

8. Start the configuration of the Content Manager by executing ./cogconfig.sh from an X Windows session.

> **Note:** When the Cognos Configuration routine starts, it creates a lock file called cogstartup.lock in <cognos installation path>/configuration directory. If your terminal session ends unexpectedly while you are running Cognos Configuration, the cogstartup.lock file might not be removed automatically. In that case, you must manually remove the lock file before you can restart Cognos Configuration successfully.

9. Because we installed the content store database on DB2 rather than using the Cognos Content Database, delete the default database resource. In the Explorer window, select **Data Access** → **Content Manager**, right-click **Content Store**, and click **Delete.**

10. Configure connectivity to the Content Manager database. Right-click **Content Manager**, and then click **New resource, Database**. We entered Content Store as the name and DB2 database as the type.

11. In the properties window, we configured the values shown in Figure 10-4 on page 179.
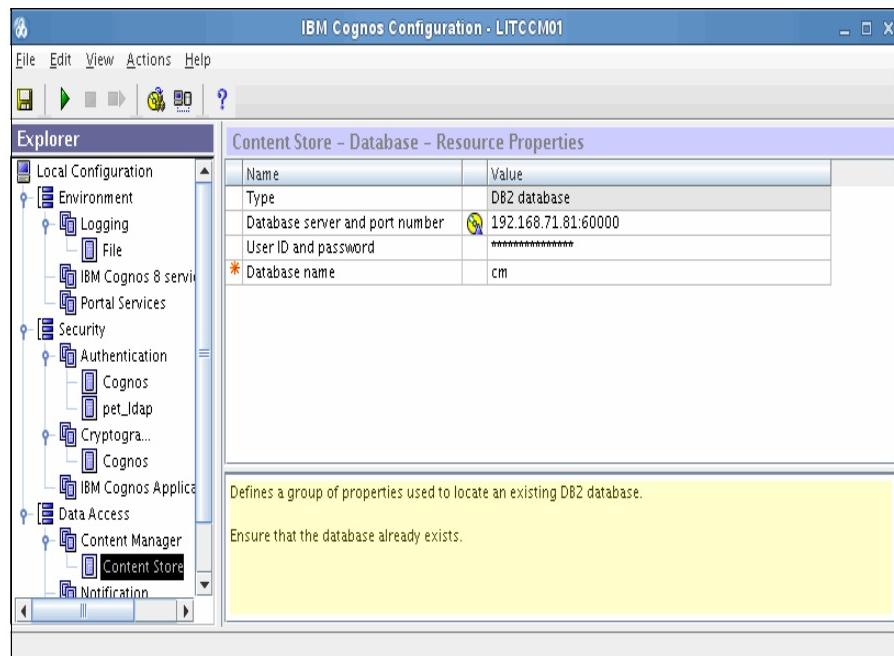
Figure 10-4   Configuring Content Manager database connection

12.From the File menu, click **Save**.

13.In the Explorer window, click **Environment**. In the Properties window, click
the value for Content Manager URIs, and then click **Edit**. In the Value -
Content Manager URIs box, click **Add**. The URIs include the IP address, port
number, and context root used by the Content Manager. The port number and
context root are populated with default values that can be changed if desired.
The first entry in the list identifies the local Content Manager and is required.
In a blank row in the table, add the full URI for the second Content Manager
computer as in Figure 10-5 on page 180. Repeat on a new blank line if
additional Content Managers are needed.
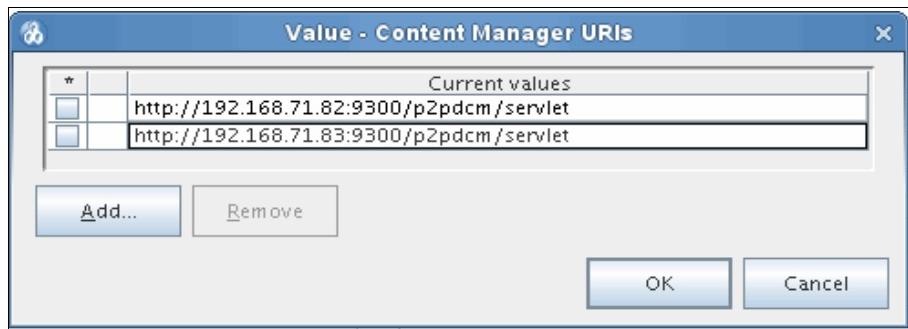
Figure 10-5   Configuring Content Manager URIs

14. In the Explorer view, go to Security and click **Cryptography**. If this is the first Content Manager to be installed, in the Properties window under CSK settings, set Store symmetric key locally to **True**. From the File menu, click **Save**.

15. Under Certificate Authority settings, set the Password property. This password must be the same for all Cognos components in the installation.

16. Under Data Access, Notification we left the entry for SMTP mail server blank because we did not want our test system to send emails.

17. Catalog the database in the **Notification** → **Notification Store**, as shown in Figure 10-6 on page 181. All Content Manager and Application Tier components must use the same notification database.
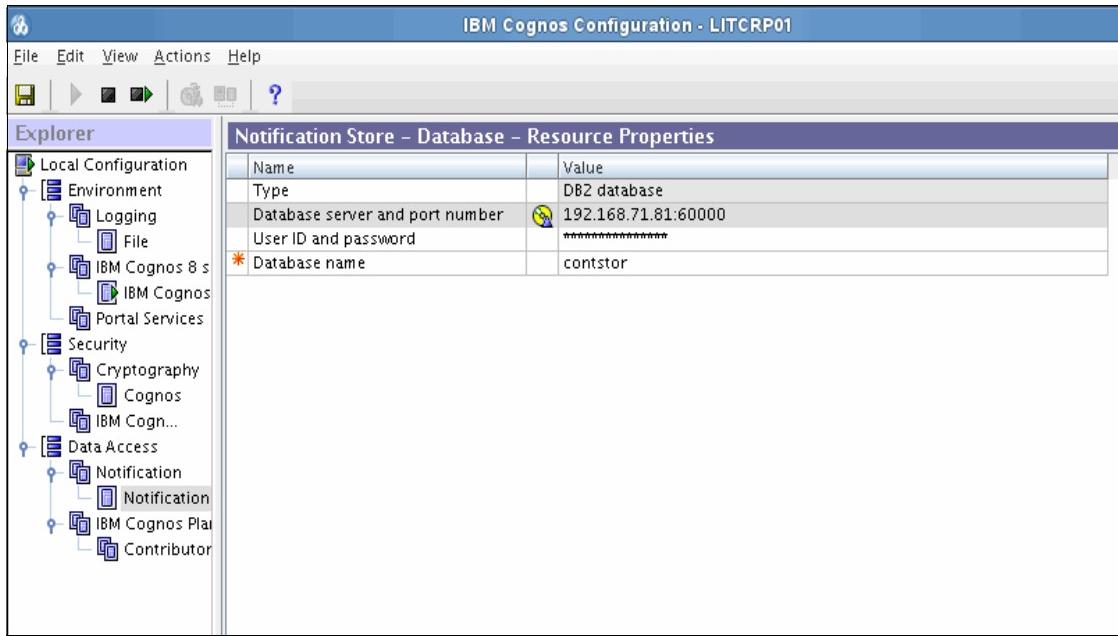
Figure 10-6   Configuration of Notification Store

18. Save the configuration changes, and start Cognos by clicking **Start**.

10.2.2 Installing and configuring the Application Tier components

We installed Application Tier components on LITCRP01 and LITCRP02. To install and configure the Application Tier components:

1. Open a terminal session through X Windows on the server where the Application Tier component is to be installed. Go to the directory where the Cognos installation code resides and start the Cognos setup program with ./issetup.

2. Follow the steps of the installation wizard.

3. In the Installation Location window, select the installation directory, as shown in Figure 10-7 on page 182.
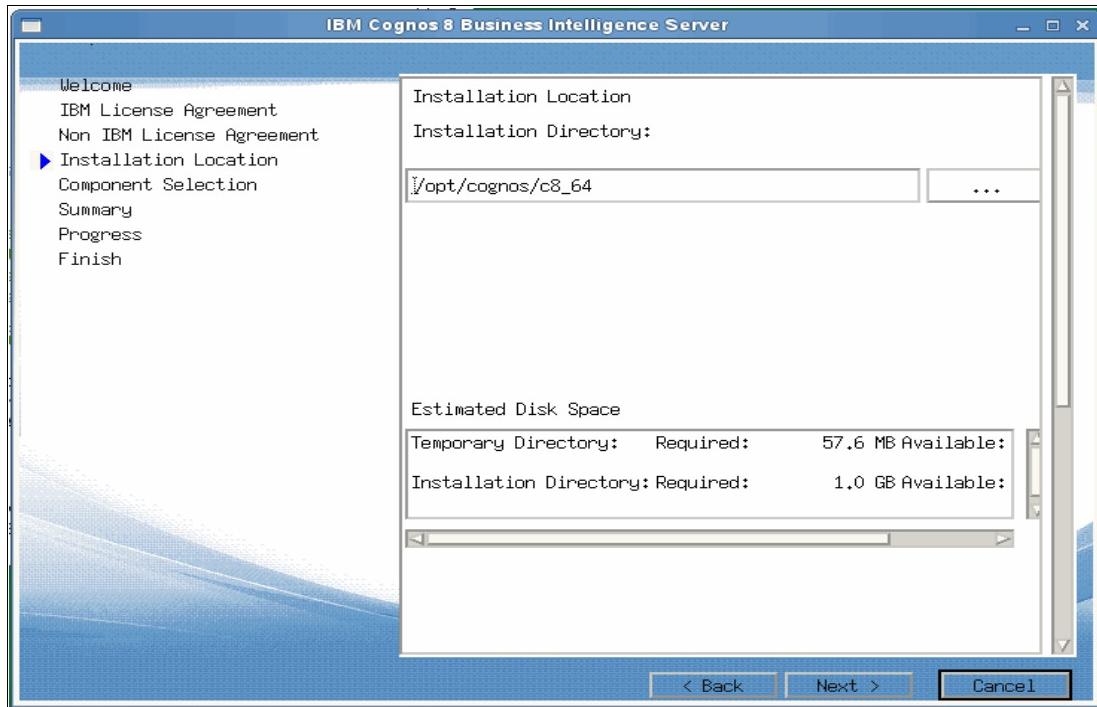
Figure 10-7   Cognos Installation Path

4. In the Component selection window, select only the Application Tier components, as shown in Figure 10-8 on page 183.
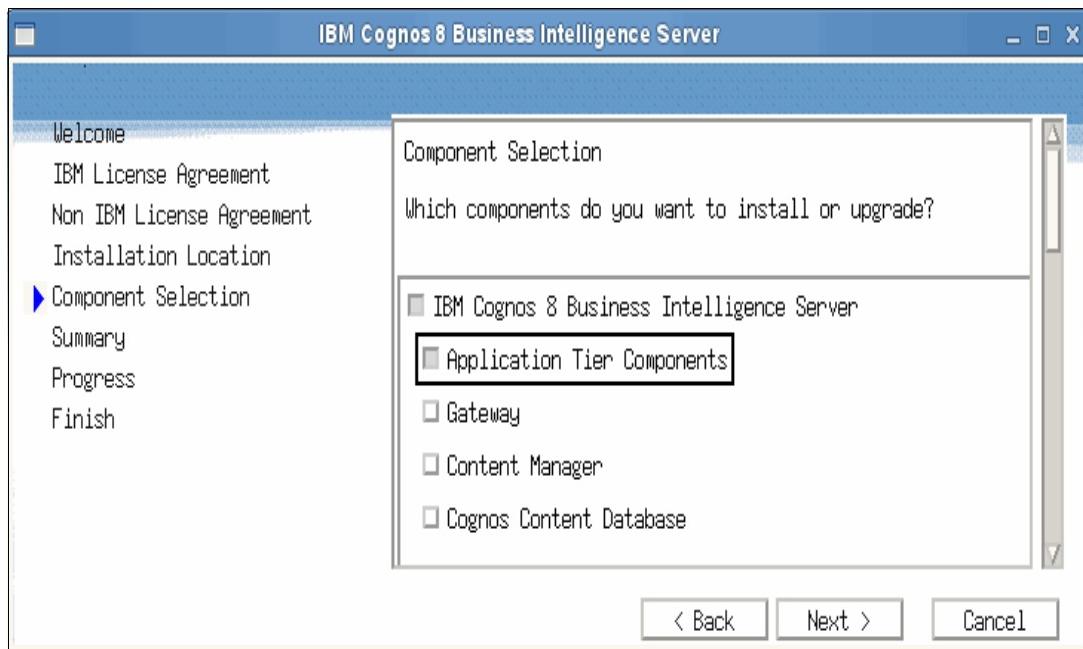
Figure 10-8   Cognos Installation Component Selection

5. The summary shows all selected components and paths, as shown in Figure 10-9 on page 184. Click **Next** to start the installation.
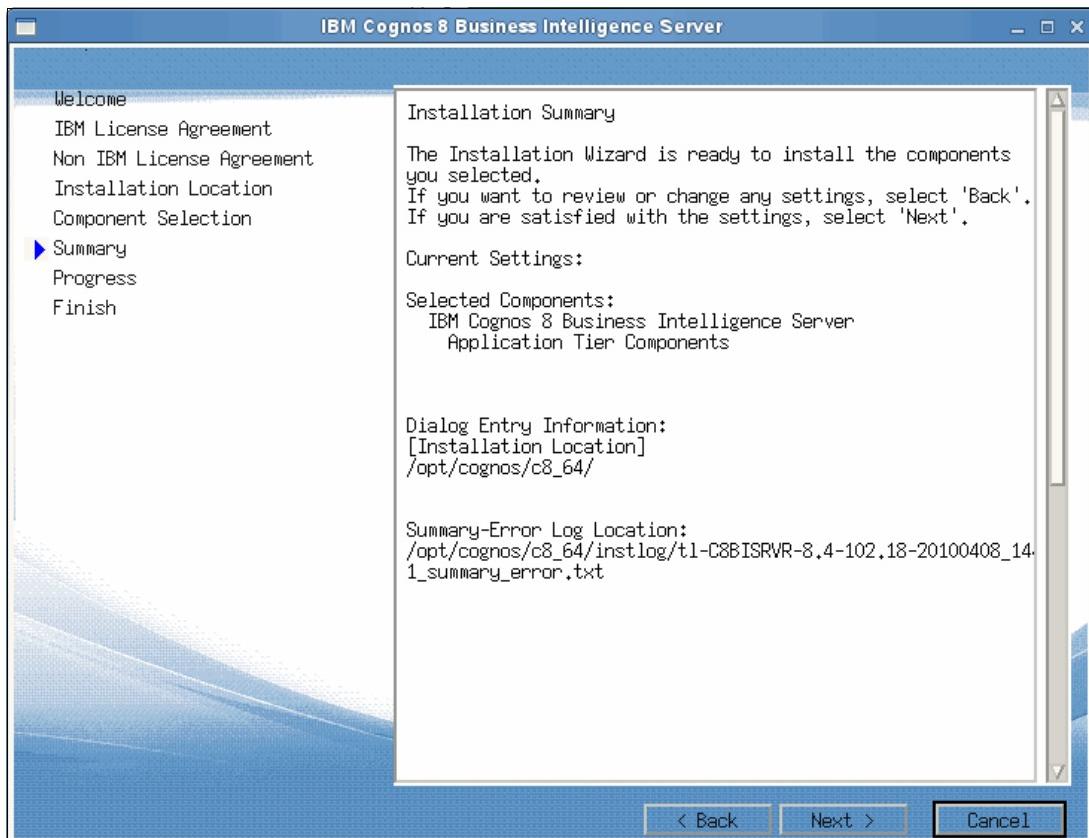
Figure 10-9   Cognos Installation Summary

6. After installation, update the LD_LIBRARY_PATH environment variable to include the appropriate library paths for the Cognos and DB2 versions installed. In our installation, we set this value in the .profile of the ID running Cognos, as shown in Example 10-1 on page 177.

7. Copy the DB2 JDBC drivers from the DB2 client installation to the Cognos installation. In our installation, we copied db2jcc.jar, db2jcc4.jar, db2java.jar and db2jcc_license_cu.jar files from /opt/ibm/db2/V9.5/java to /opt/cognos/c8_64/webapps/p2pd/WEB-INF/lib.

8. Ensure that JAVA_HOME environment variables is set to the Java Runtime Environment (JRE) location for your installation, as shown in Example 10-2 on page 178.

9. Copy the Cognos 8 cryptography file to the JRE. We copied bcprov-jdk14-134.jar from /opt/cognos/c8_64/bin64/jre/1.5.0/lib/ext to /usr/lib64/jvm/jre-1.6.0-ibm/lib/ext/.

10.Start the configuration of the Application Tier components by executing ./cogconfig.sh from an X Windows session.

11.In the Explorer pane, select **Environment**, and then change the local host portion of all URI properties to the name or IP address of the appropriate server in the Cognos installation, as show in Figure 10-10. The port numbers and context roots are populated with defaults that can be changed if desired. The Gateway URI for a Cognos Gateway. The External dispatcher URI is the HTTP endpoint through which the dispatcher receives requests from dispatchers or services on other computers. The Internal dispatcher URI is the through which the dispatcher receives requests from the local computer. The Content Manager URIs must include the URI of each Content Manager in your installation.
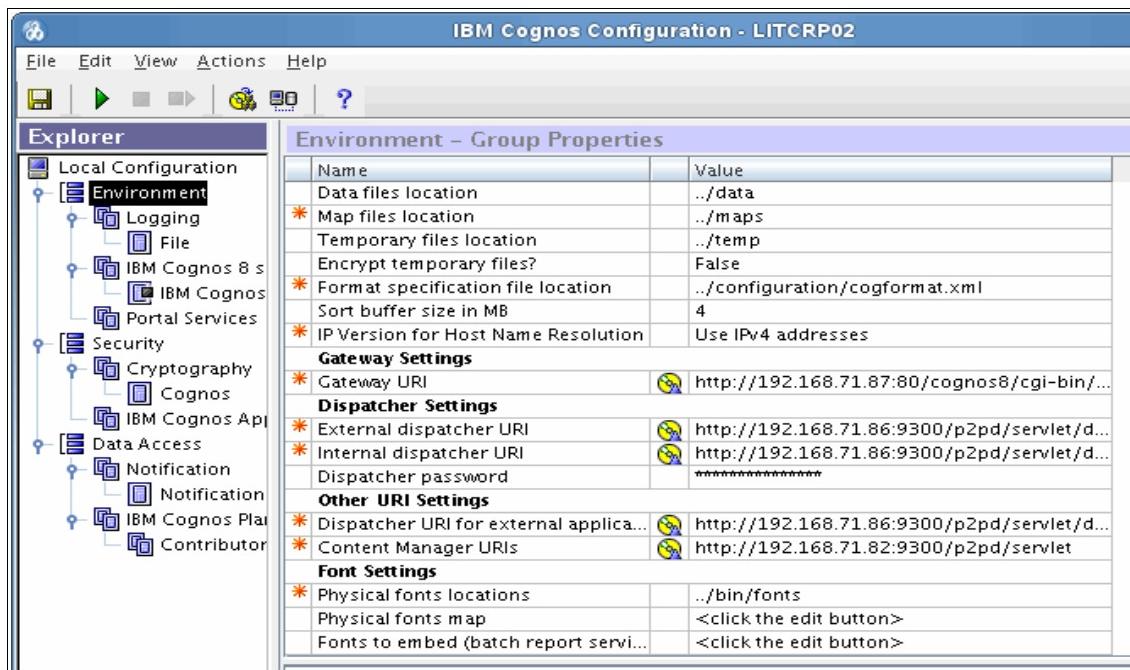
Figure 10-10   Sample Application Tier Environment configuration.

12.Select **Security** → **Cryptography** → **Cognos** in the **Explorer** pane. Under Certificate Authority Settings, set a password. This password must be the same for all the Cognos servers.

13.In the Explorer pane, select **Data Access** → **Notification** → **Notification**, catalog the database, as shown in Figure 10-6 on page 181.

14.Save the configuration by selecting **File** → **Save**. Start the Cognos Application Tier by clicking **Start**.

10.2.3 Installing and configuring the Cognos sample data

The data sources used to populate Cognos reports are not part of the Smart Analytics Cloud; however, the cloud environment must be configured to make those data sources available to users. In our lab environment, we installed the Cognos sample data as our data source on the LITCDB2 server using the sample database installation instructions in the *IBM Cognos 8 Business Intelligence Installation and Configuration Guide for Linux on System z 8.4.1*. We named the database GS_DB. We cataloged the GS_DB in the DB2 clients on LITCRP01 and LITCRP02 to provide connectivity from the reporting server to the sample database. Additional data sources can be cataloged as needed.

10.2.4 Installing and configuring the Cognos Gateway

We installed Cognos Gateway components on LITCIHS1 and LITCIHS2. To install and configure the Cognos Gateway:

1. Open a terminal session through X Windows on the server where the Application Tier component is to be installed. Go to the directory where the Cognos installation code resides and start the Cognos setup program with ./issetup.

2. Follow the steps of the installation wizard. In the Component selection window, select Gateway, and click **Next**.

3. After installing the Gateway, set the LD_LIBRARY_PATH environment variable for the ID that will run Cognos, as shown in Example 10-1 on page 177.

4. Start the configuration of the Gateway component by executing ./cogconfig.sh from an X Windows session.

5. In the Explorer pane, go to **Environment** → **Properties**, and change the Dispatcher URIs for gateway to point it to the dispatchers on the Application Tier servers, as shown in Figure 10-11 on page 187.
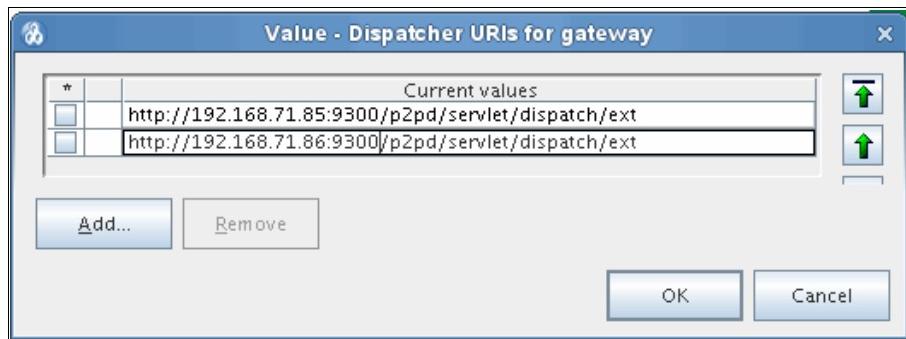
Figure 10-11   Dispatcher URI configuration for gateway.

6. To enable cryptography in the Explorer window, select **Security** →
   Cryptography, and click **Cognos** as the default cryptographic provider.
   Under the Certificate Authority settings property group, set the Password
   property to match with the password that was chosen on the default active
   Content Manager computer.

7. Test that the symmetric key can be retrieved. In the Explorer window,
   right-click **Cryptography**, and click **Test**.

8. Configure the Web server by setting up Web aliases. For Apache-based Web
   servers like IBM HTTP Server, the cognos8/cgi-bin alias must be defined
   before the cognos8 alias in the httpd.conf file as shown in Example 10-3.

Example 10-3   Sample Web aliases added to the httpd.conf file

```
ScriptAlias /cognos8/cgi-bin "/opt/cognos/c8_64/cgi-bin"
Alias /cognos8 "/opt/cognos/c8_64/webcontent"
<Directory "/opt/cognos/c8_64/cgi-bin">
        Options Indexes MultiViews
        AllowOverride None
        Order allow,deny
        Allow from all
</Directory>

<Directory "/opt/cognos/c8_64/webcontent">
        Options Indexes MultiViews
        AllowOverride None
        Order allow,deny
        Allow from all
</Directory>
```

9. Restart the Web server to make the changes available.

10. Test the Gateway. We opened a Web browser and entered the following URL for our installation: http://192.168.71.87/cognos8.

The Cognos Connection starts the web portal for IBM Cognos BI 8.4.1. as shown in Figure 10-12.
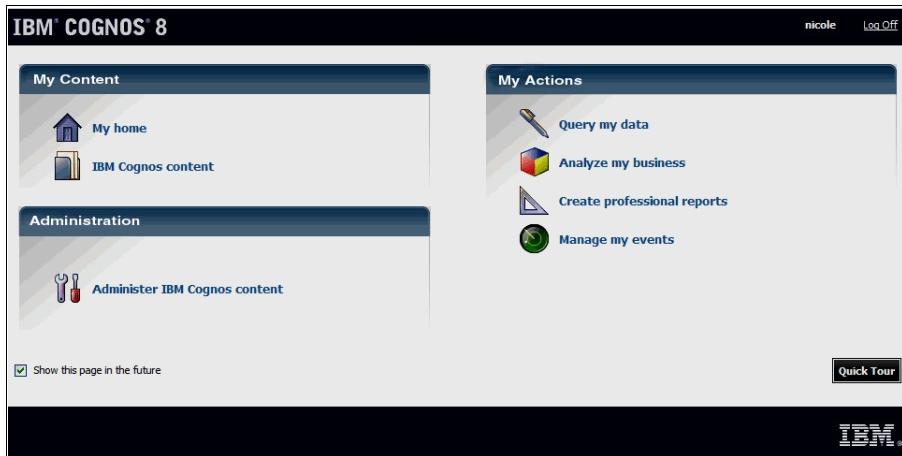
Figure 10-12   Cognos Welcome window

10.3 Connecting Cognos to the LDAP server

We configured a Cognos namespace to use the IBM Directory Server for LDAP for authentication, as shown in Figure 10-13 on page 189. For more information about the LDAP configuration, see the *IBM Cognos Administration and Security Guide*.
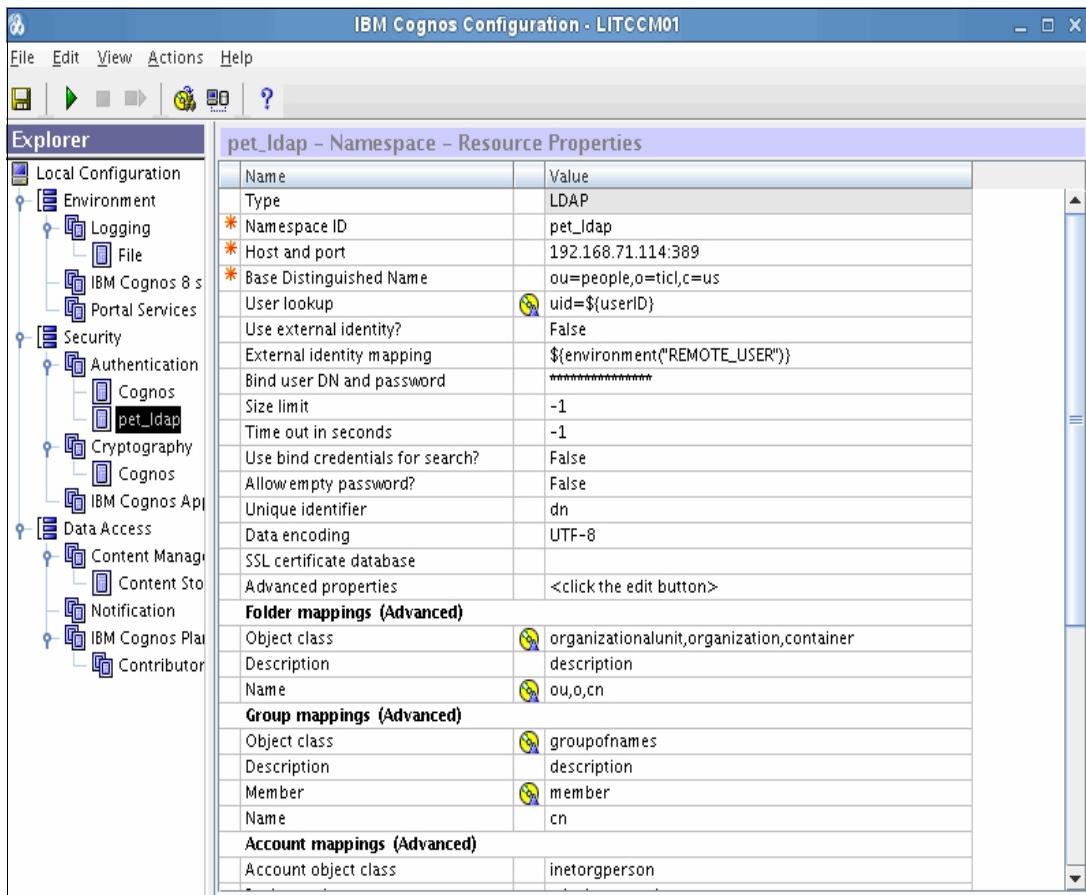
Figure 10-13   Configuring LDAP namespace within Cognos

10.4 Installing Cognos on additional servers

After the initial installation is completed successfully, the server components are distributed across multiple servers, which ensures that the Smart Analytics Cloud will become scalable and highly available.

Additional servers can be created using the Cognos installation instructions or with the provisioning tools described in Chapter 17, "Provisioning and resource management" on page 261. The Cognos servers can run as stand-alone servers at the Tomcat application server level. The dispatchers will communicate with each to provide redundancy. However, moving from Tomcat to other application servers can provide additional benefits, such as clustering, load balancing, and

centralized management. We migrated our installation to IBM WebSphere Application server for these reasons. We used WebSphere clusters to gain additional benefits, although we can run on stand-alone WebSphere servers too.

10.5 Configuring the IBM Cognos 8 to run with IBM WebSphere Application Server

IBM Cognos 8 installs with Tomcat as the application server by default. You must complete the initial installation and configuration with Tomcat before you can move to another application server like WebSphere Application Server. After building and testing our environment with Cognos running on Tomcat on servers LITCCM01, LITCCM02, LITCRP01, LITCRP02, LITCIHS1, and LITCIHS2, we chose to configure Cognos to run within IBM WebSphere Application Server. Some of the advantages of using IBM WebSphere Application Server include centralized administration for managing servers and using the WebSphere plug-in with the IBM Cognos gateway to load-balance requests to Cognos dispatchers.

The IBM WebSphere Application Server must be installed on every server with a Cognos component. Figure 10-14 on page 191 shows the key Cognos and WebSphere elements in our lab environment.
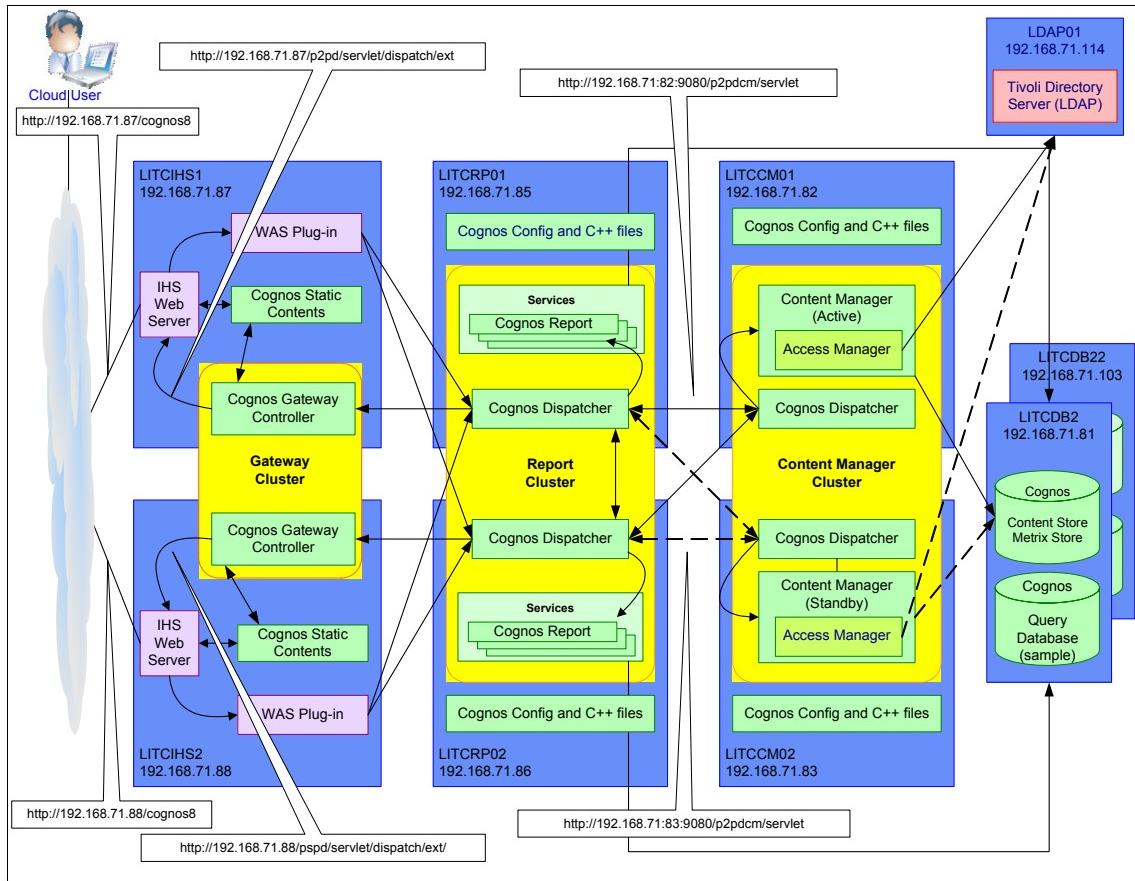
Figure 10-14   Cognos environment on WebSphere Application Server

The following sections describe the migration process used in our lab environment. Detailed instructions can be found in the *IBM Cognos 8 Business Intelligence Installation and Configuration Guide for Linux on* System z 8.4.1.

10.5.1 Backing up the current installation and setting up required environment variables

To back up the current installation and set up the required environment variables:

1. Back up the content store, as described in the *IBM Cognos 8 Administration and Security Guide*.

2. Save IBM Cognos BI 8.4.1 configuration settings in a decrypted format. From an X Windows session, start Cognos Configuration by executing

./cogconfig.sh. Backup the existing IBM Cognos information by selecting **File → Export As**.

3. Stop IBM Cognos service on the server by clicking the **Stop** button, and then close the configuration window.

4. Backup the cryptographic keys by copying them to a safe location. The files are:

```
/opt/cognos/c8_64/configuration/cogstartup.xml
/opt/cognos/c8_64/configuration/caSerial.xml
/opt/cognos/c8_64/configuration/cogconfig.prefs
/opt/cognos/c8_64/configuration/coglocale.xml
```

> **Note:** Some of the files or directories might not exist depending upon the components that are configured on the server.

Also, back up the following directories to a safe location:

```
/opt/cognos/c8_64/configuration/csk
/opt/cognos/c8_64/configuration/encryptkeypair
/opt/cognos/c8_64/configuration/signkeypair
```

5. Delete the files caSerial.xml and cogcongif.prefs, and remove the three directories: csk, encryptkeypair, and signkeypair.

6. Replace the following file with the decrypted file that was exported from the IBM Cognos Configuration. The file must use the name cogstartup.xml.

```
/opt/cognos/c8_64/configuration/cogstartup.xml
```

7. Update the Java environment:

   a. Ensure that in the .profile the JAVA_HOME environment is set to the JRE location for IBM WebSphere Application Server using

   ```
   export JAVA_HOME=/opt/IBM/WebSphere/AppServer/java/jre
   ```

   The .profile belongs to the ID running IBM WebSphere Application Server.

   b. Copy Cognos cryptography file bcprov-jdk14-134.jar from /opt/cognos/c8_64/bin64/jre/1.5.0/lib/ext/ directory file to /opt/IBM/WebSphere/AppServer/java/jre/lib/ext/.

8. Repeat steps 2-7 on all Guests where an IBM Cognos component is installed.

10.5.2 Exporting application files and configuring IBM Cognos

After all backup steps are completed and the environment variables are set, Cognos 8 must be configured with the application server information. The application is built for the IBM WebSphere Application Server. We performed the

following steps on each Content Manager, Application Tier, and Gateway server to migrate it to WebSphere Application Server:

1. Start the ./cogconfig.sh.

2. Click **Actions** → **Build Application Files**, and select the Application and Application Server Type options, as shown in Figure 10-15. In a distributed installation where there are multiple identical instances of the same type of server, for example two Application Tier servers, the application file only needs to be built one time because WebSphere Application Server can use the same file for all servers in the cluster.
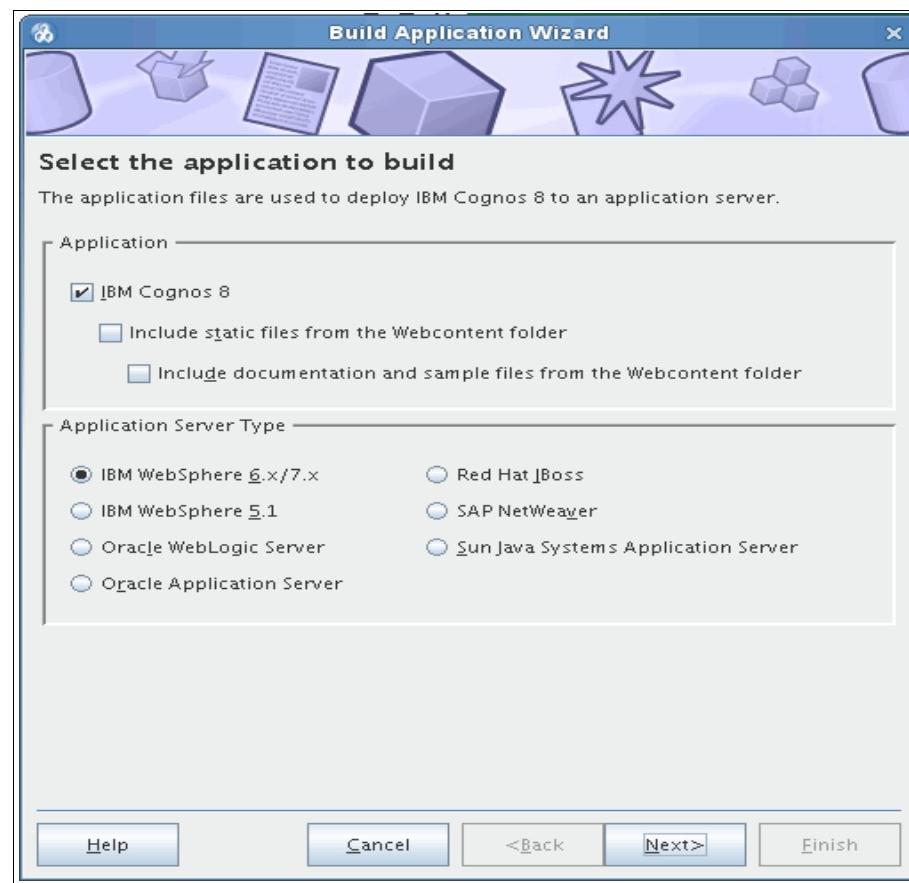
Figure 10-15   Sample Build Application Wizard

3. Click **Next**. Select the type of file to build (EAR in our case), the filename, such as p2pd.ear, location to save it, and the Context Root name. Click **Next** to build the file.

4. Determine the IBM WebSphere Application Server ports for WC_defaulthost in the IBM WebSphere Application Server Administration console. This port number is needed to update the Cognos configuration:

   a. Open the IBM WebSphere Application Server Administration Console in the web browser. Select the appropriate application server that is installed on the same guest as the Cognos component.

   b. Open the port's specification.

   c. Select the port number specified in the WC_defaulthost section, as shown in Figure 10-16.

☐	WC_defaulthost	*	9080

Figure 10-16   Port for the WC_defaulthost

5. Change the IBM Cognos configuration to use the IBM WebSphere Application Server settings. Open the configuration application ./cogconfig. In the Explorer window, select **Environment**:

   – Edit the following URIs as needed: Dispatcher URIs for Gateway, External dispatcher URI, Internal dispatcher URI, Dispatcher URI for external applications, Gateway URI, and Content Manager URIs.

   – Change the port number in the URI to the one used by IBM WebSphere Application Server as shown in Figure 10-17 on page 195. If you change the context root from the default value of p2pd change the context root in the URI.

   – On the gateway servers, we configured the Dispatcher URIs for gateway to use the WebSphere Application Server plug-in on the web server to load balance requests to the dispatchers on the Application Tier, as shown in Figure 10-18 on page 195. The WebSphere Application Server plug-in simplifies server management in a dynamic cloud environment because it adjusts load balancing to dispatchers automatically as Application Tier servers are added or removed without requiring additional configuration of the gateway within Cognos.
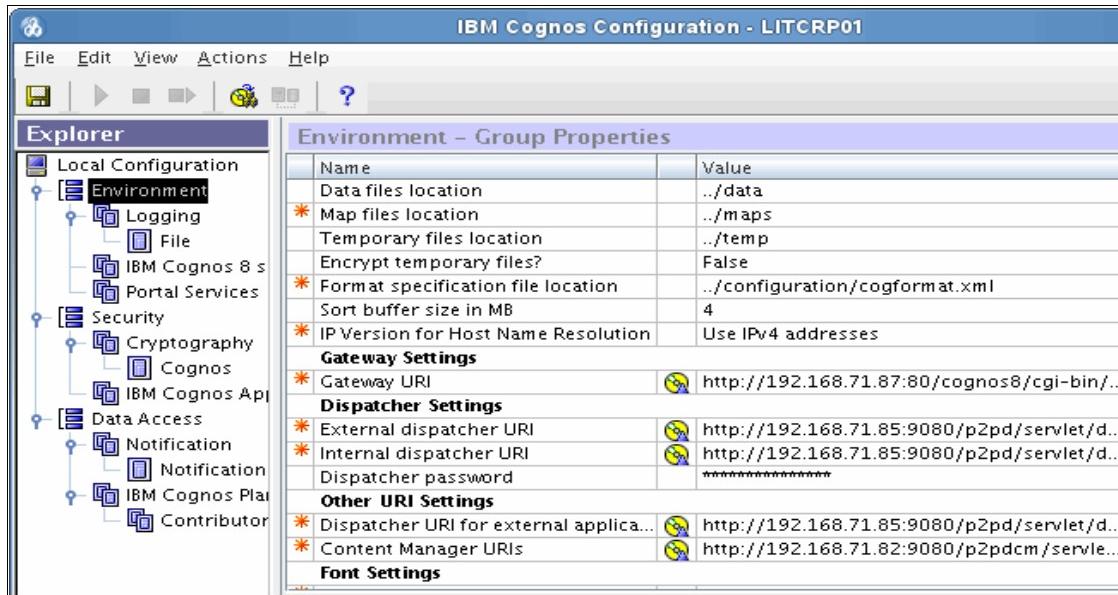
Figure 10-17   Configuring LITCRP01 URIs for WebSphere Application Server
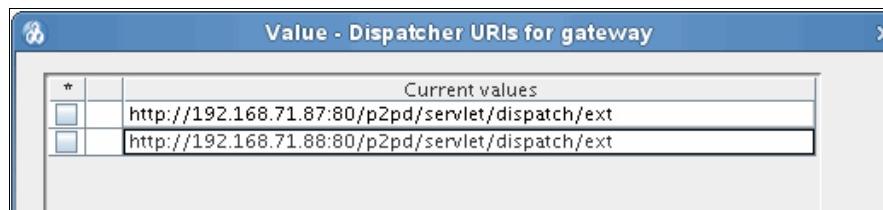
Figure 10-18   Configuring a gateway to use WebSphere Application Server plug-in

6. In the Explorer window, under Environment, IBM Cognos 8 service, right-click IBM Cognos 8, and select **Delete** to remove the Tomcat application server and eliminate the risk of starting.

7. Select **File** → **Save** to save the configuration. New cryptographic keys are generated.

8. Close the Cognos configuration.

10.5.3 Configuring WebSphere Application Server and deploying Cognos components

To complete the installation, we must update the WebSphere properties and deploy the application file:

1. Open the WebSphere Administrative console. Install a new Enterprise Application using the appropriate EAR file created in step 3 of Section 10.5.2, "Exporting application files and configuring IBM Cognos" on page 192. For example, we deployed the Application Tier using the file p2pd.ear, as shown in Figure 10-19 on page 197.

2. The default context root for both the Application Tier and Content Manager is p2pd. When deploying the Content Manager and Application Tier as separate applications within WebSphere, WebSphere Application Server does not allow two applications to have the same name or context root. During the WebSphere deployment, we assigned the name IBM Cognos 8 CM to the Content Manager with the context root p2pdcm and named the Application Tier IBM Cognos 8 with a context root of p2pd.
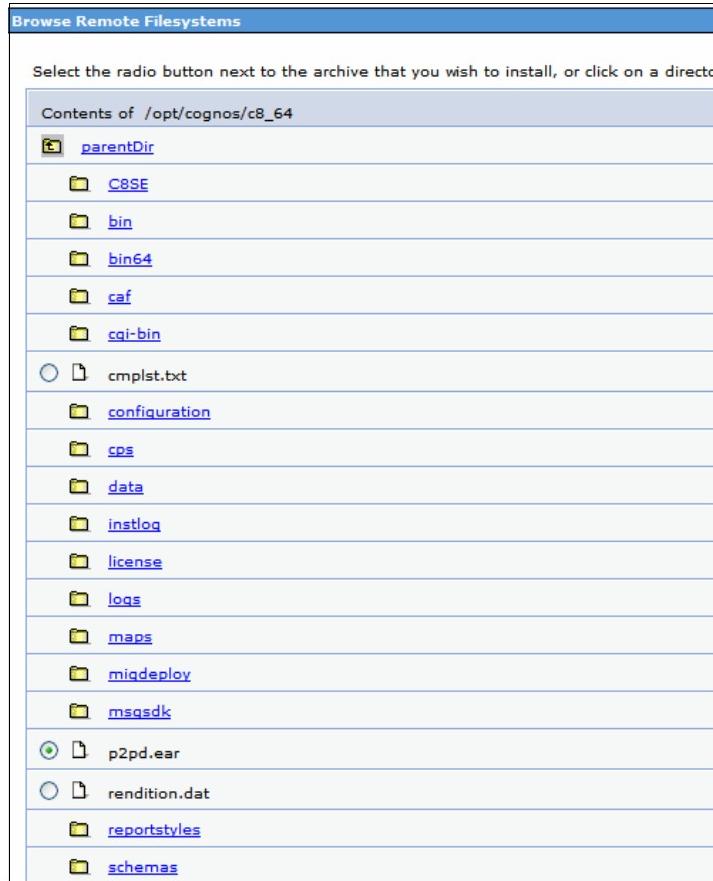
Figure 10-19   Select the appropriate .EAR file

3. Set the memory used by the JVM machine, as shown in Figure 10-20 on page 198. Set the minimum to 256 MB and the maximum to 768 MB.
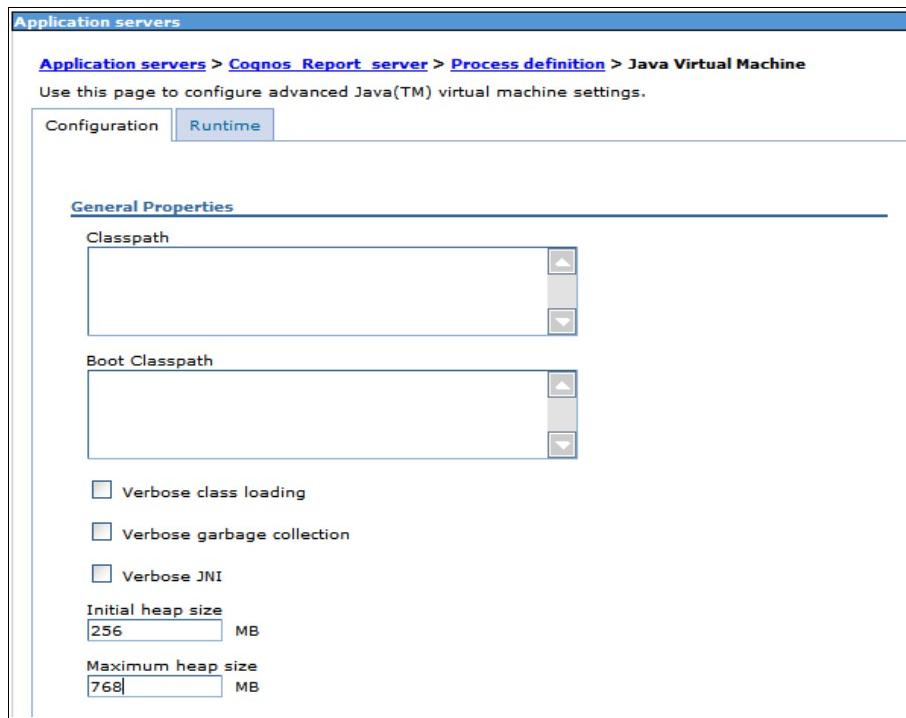
Figure 10-20   Set the memory used by JVM

4. Set the LD_LIBRARY_PATH environment variable to reference the installation_location directory in the server properties section in the IBM WebSphere Application Server administration console, as shown in Figure 10-21.

Figure 10-21   Set environment variables during WebSphere deployment

5. Stop and restart the IBM WebSphere Application Server.

6. Check the WebSphere log file SystemOut.log to verify that the server started correctly. In our lab environment, the log is in the directory /opt/IBM/WebSphere/AppServer/profiles/Appsrv01/logs/Cognos_Report_server. A successful start has log messages like those in Example 10-4.

Example 10-4   Sample log for starting Cognos on WebSphere Applicaiton Server

```
[4/23/10 17:11:50:435 EDT] 00000018 ApplicationMg A   WSVR0221I:
Application started: IBM Cognos 8
[4/23/10 17:11:50:438 EDT] 00000018 CompositionUn A   WSVR0191I:
Composition unit WebSphere:cuname=IBM Cognos 8 in BLA
WebSphere:blaname=IBM Cognos 8 started.
[4/23/10 17:11:50:458 EDT] 00000000 TCPChannel    I   TCPC0001I: TCP
Channel TCP_1 is listening on host *  (IPv6) port 9061.
[4/23/10 17:11:50:471 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain WCInboundAdmin.
[4/23/10 17:11:50:481 EDT] 00000000 TCPChannel    I   TCPC0001I: TCP
Channel TCP_2 is listening on host *  (IPv6) port 9080.
[4/23/10 17:11:50:488 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain WCInboundDefault.
[4/23/10 17:11:50:494 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain HttpQueueInboundDefault.
[4/23/10 17:11:50:497 EDT] 00000000 TCPChannel    I   TCPC0001I: TCP
Channel TCP_4 is listening on host *  (IPv6) port 9443.
[4/23/10 17:11:50:511 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain
HttpQueueInboundDefaultSecure.
[4/23/10 17:11:50:515 EDT] 00000000 TCPChannel    I   TCPC0001I: TCP
Channel TCP_3 is listening on host *  (IPv6) port 9044.
[4/23/10 17:11:50:520 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain WCInboundAdminSecure.
[4/23/10 17:11:50:526 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain WCInboundDefaultSecure.
[4/23/10 17:11:50:529 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain SOAPAcceptorChain1.
[4/23/10 17:11:50:530 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain SOAPAcceptorChain2.
[4/23/10 17:11:50:532 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain SOAPAcceptorChain3.
[4/23/10 17:11:50:534 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain SOAPAcceptorChain4.
[4/23/10 17:11:50:535 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain SOAPAcceptorChain5.
[4/23/10 17:11:50:545 EDT] 00000000 WSChannelFram A   CHFW0019I: The
Transport Channel Service has started chain SOAPAcceptorChain6.
```

```
[4/23/10 17:11:50:553 EDT] 00000018 SchedulerServ I   SCHD0077I: The
Scheduler Service is starting the Schedulers.
[4/23/10 17:11:50:556 EDT] 00000018 SchedulerServ I   SCHD0078I: The
Scheduler Service has completed starting the Schedulers.
[4/23/10 17:11:50:633 EDT] 00000000 RMIConnectorC A   ADMC0026I: The
RMI Connector is available at port 9810
[4/23/10 17:11:50:663 EDT] 00000000 JMXConnectors I   ADMC0058I: The
JMX JSR160RMI connector is available at port 9810
[4/23/10 17:11:50:682 EDT] 0000001a UserManagemen I   CWWIM6002I
Received notification that the server has finished starting.
[4/23/10 17:11:55:127 EDT] 0000001a authz         I   CWWIM2000I
Initialization of the authorization component completed
successfully.
[4/23/10 17:11:55:141 EDT] 0000001a UserManagemen I   CWWIM6003I
Initialization of the dynamic reload manager completed successfully.
[4/23/10 17:11:55:802 EDT] 00000000 WsServerImpl  A   WSVR0001I:
Server Cognos_Report_server open for e-business
```

7. Log in to the Cognos Connection web portal, and test the application
   functionality.

10.6 Using Framework Manager

Framework Manager is the Cognos tool used by administrators and developers
to create and manage metadata that serves as the basis for reports and
analyses. In our lab environment, the Framework Manager is installed on a
Windows 2003 computer.

After the installation completes, you must configure the Framework Manager. To
configure the Framework Manager:

1. Open the configuration program, as shown in Figure 10-22 on page 201.

Figure 10-22   Open Framework Manager Configuration

2. Change the Gateway URI to
   http://192.168.71.87:80/cognos8/cgi-bin/cognos.cgi, and adjust the
   Dispatcher URI for external applications to the Report Server, which in our lab
   environment is http://192.168.71.85:9080/p2pd/servlet/dispatch, as shown in
   Figure 10-23 on page 202.
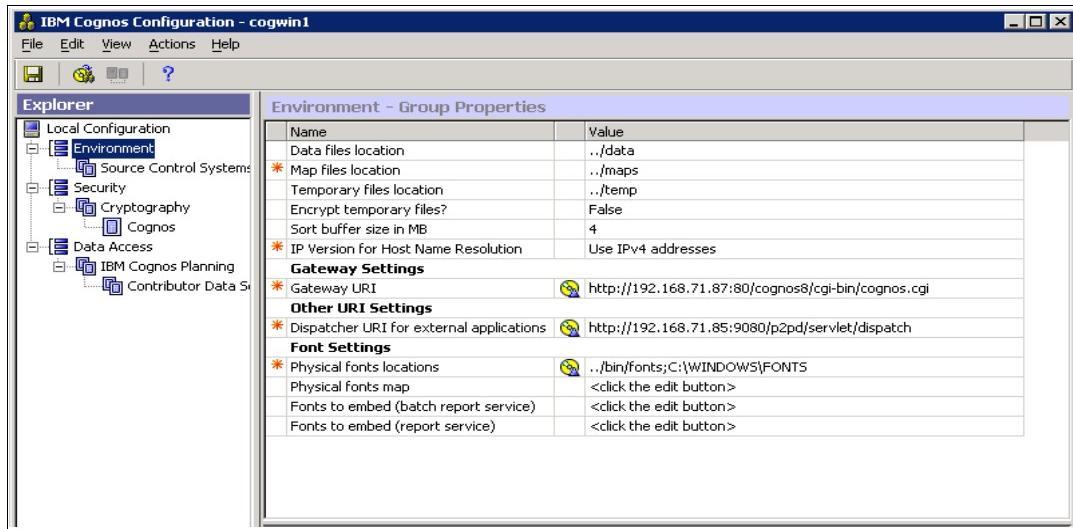
Figure 10-23   Framework Manager Configuration settings

After the configuration is complete the Framework Manager can be started.

10.6.1 Cataloging a new data source

When a new database must be made available in the Framework Manager, it must be registered. Make sure the database is cataloged on the server that is indicated in the Framework Manager Configuration Dispatcher URI for external applications.

1. When prompted for a session command in XML format use:

```
<commandBlock>
<commands>
<sqlCommand>
<sql>SET CURRENT SCHEMA = GOSLDW</sql>
</sqlCommand>
</commands>
</commandBlock>
```

2. Click **OK**. An excerpt of the command is displayed in the value column.

3. Click **Finish** to complete creating the data source.[1]

---

[1] http://www.clearviewinformatics.com/2008/04/23/setting-up-the-cognos-8-samples-on-db2/

Infrastructure security

Business information is strategic for the enterprise and has therefore tight requirements on security. When providing information in a private cloud these requirements have even a stronger focus. A consistent role and rights management throughout the company can best be implemented centralized.

This chapter provides an example of implementing authentication in Lightweight Directory Access Protocol (LDAP) for Linux systems and securing z/VM with Resource Access Control Facility (RACF).

Information about the security features of the particular software components that we use in the Smart Analytics Cloud is in the chapters where the implementation of the particular component is described.

> **Note:** In this IBM Redbooks publication, we cover basic security features and give an example of their implementation on SLES 11. You can find more about Security for Linux on System z in *Security for Linux on System z*, SG24-7728.

11.1 Centralization of Linux security with LDAP

These days most software vendors provide the possibility to authenticate in their products using LDAP. This process of authentication allows organizations to avoid defining users for each software component and to use unique user names and passwords for all services instead.

IBM Tivoli Directory Server (ITDS) is an IBM implementation of LDAP server that is available for almost every platform (for example, Linux, AIX, Windows, Solaris, z/OS, z/VM, and so on). For our installation, we use the existing ITDS version 6.1 on Linux for System z because we assume that the user of our Smart Analytics Cloud Solution will integrate it in a current IT infrastructure and will not build it from scratch.

There are two subsystems in Linux that must be configured to make the possible authentication in LDAP:

▶ Plugable Authentication Modules (PAM)
▶ Name Service Switch (NSS)

PAM handles authentication tasks of most applications on the system. It provides the ability for the system administrator to choose separate mechanisms for accounting, authentication, password policies, and session settings for various applications.

NSS handles calls to a function that retrieves data from a system database, including password and group information.

11.1.1 Enabling LDAP authentication with YaST

To enable LDAP authentication, PAM and NSS configuration must be changed properly. You can manually edit the files in directory /etc/pam.d, /etc/ldap.conf and /etc/nsswitch.conf. However, it is more comfortable to use YaST for this task:

1. Run `yast` from the command line, and select LDAP Client from the Network Services menu, as shown in Figure 11-1 on page 205.

```
YaST2 - menu @ LITCDB2

                          YaST2 Control Center


  Software                  DHCP Server
  Hardware                  DNS Server
  System                    FTP Server
  Network Devices           HTTP Server
  Network Services          Hostnames
  Novell AppArmor           Kerberos Client
  Security and Users        Kerberos Server
  Support                   LDAP Browser
  Miscellaneous             LDAP Client
                            LDAP Server
                            Mail Transfer Agent
                            NFS Client
                            NFS Server
                            NIS Client
                            NIS Server
                            NTP Configuration
                            Network Services (xinetd)
                            Proxy
                            Remote Administration (VNC)
                            SLP Server
                            SSHD Configuration
                            Samba Server
                            Squid
                            TFTP Server
                            WOL
                            Windows Domain Membership
                            iSCSI Initiator
                            iSCSI Target
                            iSNS Server
```

Figure 11-1   Selecting LDAP configuration in YaST

2. In the User Authentication section, select the Use LDAP option. Type `LDAP hosts` and `Base Distinguished Name`, as depicted in Figure 11-2 on page 206.

```
YaST2 - ldap @ LITCDB2

 LDAP Client Configuration




    ┌User Authentication────────────────────────────────────────────────────────────┐
    │ ( ) Do Not Use LDAP                                                            │
    │ (x) Use LDAP                                                                   │
    │ ( ) Use LDAP but Disable Logins                                               │
    └───────────────────────────────────────────────────────────────────────────────┘
    ┌LDAP Client────────────────────────────────────────────────────────────────────┐
    │ Addresses of LDAP Servers                                                     │
    │ ldap01.ltic.pok.ibm.com ldap02.ltic.pok.ibm.com░░░░░░░░░░░░░░░░░░░░░░░░[Find]  │
    │ LDAP Base DN                                                                   │
    │ ou=People,o=ticl,c=us░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░[Fetch DN]     │
    │ [ ] LDAP TLS/SSL                                                              │
    │ [ ] LDAP Version 2                                                            │
    └───────────────────────────────────────────────────────────────────────────────┘

    [x] Start Automounter
    [ ] Create Home Directory on Login
                                        [Advanced Configuration...]




 [ Help ]                            [Cancel]                                  [  OK  ]
 F1 Help  F8 Cancel  F10 OK
```

Figure 11-2   Filling in LDAP configuration in YaST

3. Click **Advanced configuration**, and change Naming Context, Password
   Change Protocol, and Group Member attribute, if necessary, as shown in
   Figure 11-3 on page 207.

```
YaST2 - ldap @ LITCDB2

Advanced Configuration
 ┌Client Settings──Administration Settings─────────────────────────────────────────────
 │ ┌Naming Contexts─────────────────────────────────────────────────────────────────
 │ │  User Map
 │ │  ou=People,o=ticl,c=us▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓[Browse]
 │ │  Password Map
 │ │  ou=People,o=ticl,c=us▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓[Browse]
 │ │  Group Map
 │ │  ou=People,o=ticl,c=us▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓[Browse]
 │ │
 │ │  Password Change Protocol
 │ │  crypt▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓↓
 │ │
 │ │  Group Member Attribute
 │ │  member▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓↓








[ Help ]                              [Cancel]                                    [  OK  ]
 F1 Help  F8 Cancel  F10 OK
```
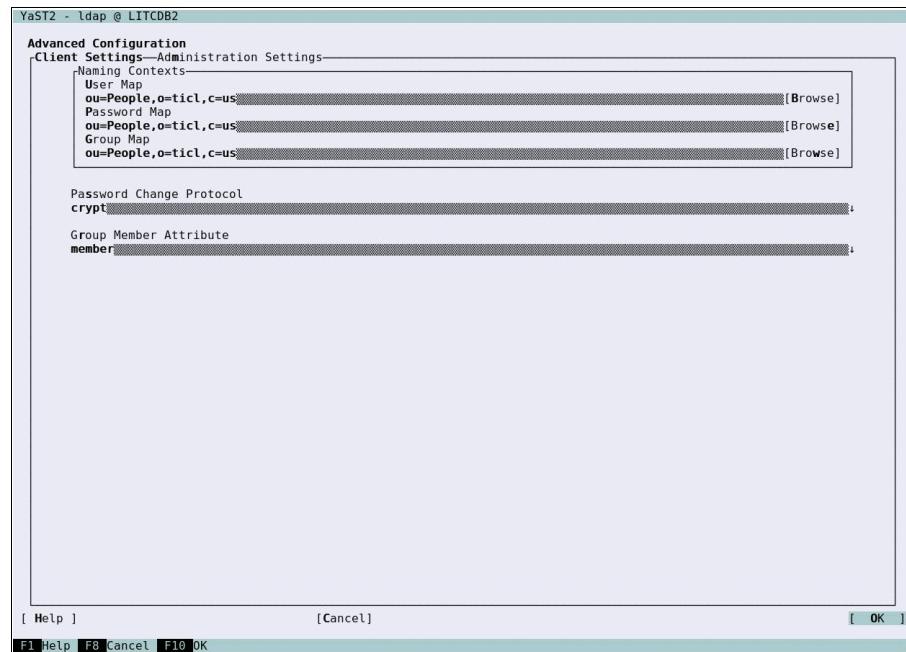
Figure 11-3   Advanced configuration of LDAP in YaST

11.1.2 Troubleshooting the LDAP connection

If the LDAP authentication does not work, the first thing to check is whether the problem is caused by the connection to the LDAP server or by the configuration of settings.

To check the LDAP connection, the openldap2-client package must be installed. Using this command line utility, as shown in Example 11-1, you can connect to the LDAP server.

Example 11-1   Checking if openldap2-client is installed

```
LITCDB2:~# rpm -q openldap2-client
openldap2-client-2.4.12-7.16
```

If the openldap2-client utilities are installed, can try to connect to your LDAP server as shown in Example 11-2 on page 208.

Example 11-2   Checking connection to LDAP server

```
LITCDB2:~# ldapsearch -x -h ldap01 -p 389 -D "cn=root" -w password -b
"" "objectclass=*"
```

If the connection to the LDAP server works correctly, you will see the LDAP tree. If not, perform the same operations from the server itself to check if the issue is related to a networking problem or other problems with ITDS.

The next step is to check if the NSS subsystem works correctly. For this purpose you can use the `getent` command, as shown in example Example 11-3.

Example 11-3   Checking Name Service Switch

```
LITCDB2:~# getent passwd
```

If it does not show LDAP users, check /etc/ldap.conf, if you used the right values for the main parameters, such as LDAP server, distinguished name, object class. and so on.

11.2 Improving z/VM security with RACF

z/VM, itself, provides a set of built-in functions to make guest machines independent from each other, introduce concepts of privileges for users, and secure internal networks. To make available managing z/VM security from one point and improve audit features use External Security Manager (ESM).

Resource Access Control facility (RACF) is an ESM that comes as an optional feature of z/VM distribution. It performs the following operations:

► Identifies the users that connect to the system and checks their identity.
► Gives these users access to the system resources under its surveillance.
► Records and reports access to the system.

Refer to the *Program Directory for RACF Security Server*, GI11-2894 for information about step-by-step implementation of RACF.

11.3 z/VM LDAP Server and RACF

At the beginning of this chapter, we mentioned that we use a preconfigured IBM Tivoli Directory Server on Linux for System z. If no LDAP server is available when implementing a System Analytics Cloud, it is possible to use the z/VM LDAP Server for this purpose. It can be integrated with RACF so that not only is

authentication of Linux guests and software centralized but also z/VM. For further details about this topic, see *Security on z/VM*, SG24-7471-00.

z/VM has a feature called $logon\ by$ that allows the cloud service administrators to connect to z/VM with MAINT privileges using their own user ID. This same functionality can be implemented in Linux using $sudo$.

12

Onboarding application

Our onboarding application is built on top of a web forms creation tool, which handles database tables to store the data and user input to work with this data. This tool is a standard J2EE application that is deployed to the IBM WebSphere Application Server and IBM DB2 Universal Database. It uses a database as its operational data store for the development environment and to store the business data of created applications, such as the data generated from our onboarding application.

We configured our onboarding application to use an existing LDAP server as our authentication provider (see section 9.7, "Enabling security" on page 151). Within the application, you can define a finer-grained security setting.

As an alternative to this solution, the IBM Tivoli Service Request Manager can be used as a preconfigured tool that allows the selection of services from a services catalog and that provides basic workflow functionality to deliver the requested service.

13

Provisioning and resource reallocation

Automated provisioning is an important part of the Smart Analytics Cloud and one of the main differentiators from regular installation of Cognos. The System z platform provides a lot of features to simplify deployment of systems and configuration tasks, but automation is still needed in complex environments.

There are several products that are available for System z that can provide automation mechanisms for provisioning operating systems and software. We used IBM Workload Mobility Workbench—a product that you can order as an IBM Service offering. Contact your IBM representative to get more information about this product.

In this chapter, we provide an overview about the IBM Workload Mobility Workbench installation, the necessary steps for preparation of software images, and at the end we discuss preparation of systems for dynamic resource management, which enables CPU, memory, and I/O to be reallocated between systems *on the fly*.

13.1 Prerequisites for IBM Workload Mobility Workbench

IBM Workload Mobility Workbench consists of a management server that provides a web interface for the *Cloud Administrator* for creating provisioning requests and a client, which receives these requests and implements provisioning on the system. For System z, the client is an application running on z/VM.

Installation of IBM Workload Mobility Workbench can be done only by an authorized service engineer, so the purpose of this chapter is not to give a step-by-step implementation instruction for this product; instead, we aim to give an overview of the changes that occur in your environment and how to operate in it.

IBM Workload Mobility Workbench requires z/VM version 5.1 or higher.

The User Directory of z/VM—a component where information about guest machines is stored, must be managed with DirMaint™ service. DirMaint is an optional component that comes with z/VM. It provides functionality for automated operations and conflict resolutions of simultaneous requests. Refer to *Security on z/VM*, SG24-7471-00 for step-by-step installation instructions.

IBM Workload Mobility Workbench requires the archiving utility VMARC to be installed on z/VM. Refer to *Illustrated Guide to Using VMARC*, Techdoc PRS3332 about installation instructions and usage.

Security of z/VM can be optionally managed by Resource Access Control Facility (RACF), which we discuss in 11.2, "Improving z/VM security with RACF" on page 208.

13.2 Installing IBM Workload Mobility Workbench

Installation of IBM Workload Mobility Workbench requires creating special service guests on z/VM running in disconnected mode with enough authority to execute privileged commands. They receive requests from the management server and do necessary operations to make new systems available from golden images.

13.3 Preparing software images

The Provisioning Asset Library stores information about the various components that are necessary to deploy new servers. These components include target hosts, such as a target z/VM LPAR or a private cloud, and images, such as predefined servers, and software (individual software components).

For our Smart Analytics Cloud, the following servers are mandatory to be available for fast provisioning:

► Cognos reporting server (litcrp01)
► Cognos gateway server (litcihs1)

The node with the Cognos Content Manager server has only one active image. Any additional server act as a standby, and cloning the third one is optional in our lab environment. We created a post-provisioning script in case we need it.

The Deployment Manager can be cloned, but it will be used as a backup image only for recovery purposes. There is no need to run a post-provisioning script because it is brought up only when the active Deployment Manager is shutdown because of a failure.

LITCBA01 and LITCBA02 is solely for our lab environment onboarding application; therefore, there is no need for cloning. In a real scenario, it might need to be cloned for scalability reasons.

LITCDB2 has it own secondary server LITCDB22 for HADR, and there is no need for cloning. Therefore, there is no post-provisioning script required.

13.3.1 Software images considerations

There are two general ways to make a software image:

► A silent installation scenario where a vanilla image is used and then augmented with software packages at provisioning time.

► A direct clone approach where a complete image (known as a golden image) includes a preconfigured operating system, users, groups, file systems, and software.

The silent installation approach relies on being able to deploy software products using automated installers (usually provided with the product). This approach is useful in two particular scenarios:

► When the target server might have many unique variations of a particular software configuration.

► When the target servers have many variations of software products.

The downside of this approach is, that it introduces more opportunities for errors (network install dependencies, prerequisites, and so on) and generally results in longer deployment times due to a complete installation process.

In our case, the target Cognos servers are identical for each deployment. We do not expect the need to have any additional software installed or operating system customizations made. The sole change that is performed on the new server will be the modification of network related configurations (IP address, host name). This modification is performed through scripts that reconfigure the new server immediately after it is cloned from golden image.

13.3.2 Preparing a golden image for cloning

To prepare a golden image for a cloning:

1. Install the operating system with the required users, groups, and file systems and customize it to support the desired software products. This task includes steps like installing prerequisite operating system packages, kernel parameters, and drivers.

2. Install all of the required software components and test that they function as expected.

3. Modify the server for integration with the IBM Workload Mobility Workbench (customization scripts mentioned previously).

13.3.3 Taking images from software and adding to the asset library

At this point the image is loaded into the Provisioning Asset Library, which makes it available to be used within the provisioning windows. To do this, the z/VM guest ID and node and metadata must be defined, which helps users distinguish one image from another.

13.3.4 Post-provisioning scripts

Post-provisioning scripts are created to make necessary changes in network configuration of new servers. In our case, the post-provisioning scripts are for the operating system, WebSphere Application Server, and the Cognos servers.

Operating system

After the Linux image was cloned, operating system configurations, such as host name and network settings, must be modified on the cloned system. In our lab

environment, these changes occurred ahead of time as part of the operating system post-provisioning process.

WebSphere Application Server for Cognos

In our lab environment, we decided to clone only the Cognos Report server and the Cognos Gateway server. After the image is cloned, perform the following steps on the cloned image before the WebSphere Application Server becomes operational on the cloned image:

> **Note:** We assume that no WebSphere Application Server profile exist on the golden image.

1. Back up the two WebSphere Application Server properties files in the golden image:
   – soap.client.props
   – ssl.client.props

2. Re-Create a profile: Appsrv01.

3. Federate the created profile to the Deployment Manager.

4. Restore the two backed up WebSphere Application Server properties files.

5. Add the cloned image as a new member to the existing Application Cluster.

6. Update the JVM heap size and environment entries per Cognos requirements.

7. Create the web server, if it is the clone for the Cognos Gateway server.

8. Modify the httpd.conf and admin.conf for the target host name, if it is the clone for the Cognos Gateway server.

9. Start the application server.

10. Start the web server, if it is the clone for the Cognos Gateway server.

You can perform this procedure using a combination of manual editing, command line utility, and the Admin Console. In our lab environment, we created post-provisioning scripts that can automate these steps. See Appendix A, "WebSphere Application Server post-provisioning scripts" on page 289 for the sample scripts that we used after the Cognos Report server and the Cognos Gateway server were cloned.

Cognos servers

For the Cognos Reporting server we must change the dispatcher URIs in the Cognos configuration. We use a script to back up the Cognos cryptography files,

generate a decrypted cogstartup.xml with the new server information in it, and execute the Cognos configuration in silent mode.

13.4 Dynamic resource reallocation planning

System z, z/VM, and Linux provide the ability to add or reduce, *on the fly*, almost all resources. It can be useful for example for:

► Reallocating resources between the production, development, and test systems
► Utilizing resources after nondisruptive hardware upgrades
► Adding systems for new, separate task-like quality assurance

A set of changes must occur to make a system ready for *on the fly* resource management.

13.4.1 Preparing for flexible CPU reallocation

To have the ability to add or reduce CPUs for the logical partition later you must define reserved CPUs in the logical partition activation profile.

For z/VM guest virtual machines, you must define as much CPUs as you plan to use later, and then stop part of them with the `set share` command, so that z/VM will not schedule any tasks on them.

13.4.2 Preparing for dynamic memory management

The System z server provides a dynamic storage-reconfiguration capability to change the amount of operating memory available for use in a logical partition while the partition is active. The configuration of storage for LPAR includes an initial amount of memory, plus an additional reserved amount, defined in fixed-size increments. The reserved memory can be assigned and accessed on demand, using the dynamic storage-reconfiguration function.

Additional memory can come from:

► Unused available memory
► Concurrent memory upgrade
► Other LPAR released memory

Starting from version 5.4 z/VM takes advantage of this architecture for its own use and also virtualizes it for guest usage. With this capability there is no need to re-IPL (reboot) z/VM or guest systems to increase memory size.

Dynamic memory upgrade from the guest operating system side is at this moment supported just by SLES 11. You can reserve memory with the `set reserved` command for later use by a Linux guest.

13.4.3 Preparing for dynamic I/O management

To make the I/O subsystem on the machine level available for dynamic changes, you must properly configure the Hardware Configuration Definition (HCD) service in z/VM (or in z/OS if you have one). HCD is a pre-installed component of z/VM. Refer to the *HCD/HCM for z/VM Program Directory* for information.

On the Linux level, it is a good idea to place file systems, which can grow later, on LVM logical volumes.

14

Tivoli Monitoring agent implementation

This chapter focuses on the implementation of the agents necessary for the monitoring of the Smart Analytics Cloud (SAC). We show how to install and configure the agents for the various components that are deployed in the SAC.

221

14.1 Tivoli Monitoring on existing infrastructure

This book assumes that an existing IBM Tivoli Monitoring infrastructure is already implemented and operational, including the Tivoli Enterprise Management Server (TEMS) and the Tivoli Enterprise Portal Server (TEPS).

A typical IBM Tivoli Monitoring environment comprises of the following components:

► One or more Tivoli Enterprise Monitoring Servers (TEMS), which act as a collection and control point for alerts that are received from the agents and collect their performance and availability data. The monitoring server also manages the connection status of the agents. One server in each environment must be designated as the hub.

► A Tivoli Enterprise Portal Server (TEPS), which provides the core presentation layer for retrieval, manipulation, analysis, and pre-formatting of data. The portal server retrieves data from the hub monitoring server in response to user actions at the portal client and sends the data back to the portal client for presentation. The portal server also provides presentation information to the portal client so that it can render the user interface views suitably.

► One or more Tivoli Enterprise Portal clients with a Java-based user interface for viewing and monitoring your enterprise. Tivoli Enterprise Portal offers two modes of operation: desktop and browser.

► Tivoli Enterprise Monitoring Agents that are installed on the systems or subsystems that you want to monitor. These agents collect data from monitored or managed systems and distribute this information either to a monitoring server or to an SNMP Event Collector, such as IBM Tivoli Netcool/OMNIbus.

► z/OS only: Tivoli Management Services:Engine (TMS:Engine) provides common functions, such as communications, multithreaded runtime services, diagnosis (dumps), and logging (RKLVLOG), for the Tivoli Enterprise Monitoring Server, monitoring agents, and OMEGAMON components of OMEGAMON XE products running on z/OS.

► An Eclipse Help Server for presenting help for the portal and all monitoring agents for which support is installed.

An installation optionally includes the following components:

► Tivoli Data Warehouse for storing historical data collected from agents in your environment. The data warehouse is located on an IBM DB2 on the workstation, DB2 on z/OS, Oracle, or Microsoft SQL database. To store data in this database, you must install the Warehouse Proxy agent. To perform

aggregation and pruning functions on the data, you must also install the Summarization and Pruning agent.

► An event synchronization component that sends updates to situation events that are forwarded to a Tivoli Enterprise Console® event server or a Netcool/OMNIbus Object Server back to the monitoring server.

In our environment, we have one TEMS and one TEPS.

14.2 Tivoli OMEGAMON XE on z/VM and Linux

In our environment, Tivoli OMEGAMON XE on z/VM and Linux provides metrics about the z/VM hypervisor and the Linux guests to the Tivoli Monitoring infrastructure.

IBM Tivoli OMEGAMON XE on z/VM and Linux utilizes the data collection from the Performance Toolkit for VM (PTK is a prerequisite) and complements it with data collection by the ITM Linux for zSeries agent. The Performance Tool Kit is the foundation for gathering z/VM metrics. It has proven to be rich in information and provides the ideal base for z/VM data input to OMEGAMON XE on z/VM Linux. The Linux agent has been the basis for monitoring Linux on all platforms. OMEGAMON XE on z/VM and Linux takes advantage of the Tivoli Enterprise Portal (TEP) and allows for all of the Tivoli Enterprise Portal alerting, action, and integration capabilities to be utilized.

We do not describe the implementation of this piece of software in this book.

14.3 WebSphere Application Server agent implementation

In this section, we describe the installation and the configuration of the monitoring agent for WebSphere Application Server.

14.3.1 ITCAM for WebSphere with IBM Tivoli Monitoring

Figure 14-1 on page 224 displays a basic overview of how the components for IBM Tivoli Monitoring and ITCAM for WebSphere interact.
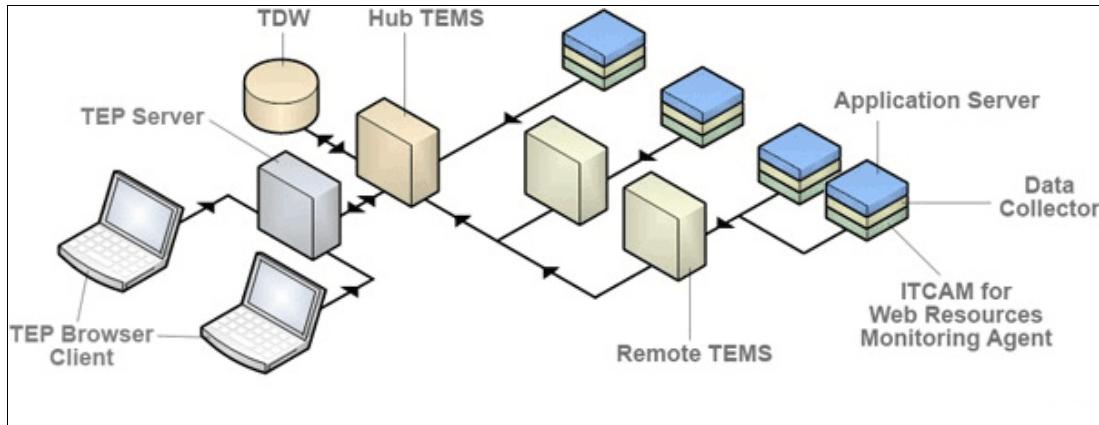
Figure 14-1   Monitoring architecture with IBM Tivoli Monitoring

The Tivoli Enterprise Monitoring Agent (TEMA) collects performance data about the WebSphere servers running on a single node from four primary sources:

► Response time data for your applications' service requests from the Data Collector

► Resource data from WebSphere's Performance Monitoring Infrastructure (PMI)

► WebSphere Application Server log messages

► Garbage-collector activity recorded in the Java virtual machine's verboseGC trace

The monitoring agent accumulates data from all of these sources. The TEMS retrieves this data, merges it with data from other monitoring agents, and passes them on to the portal server for display on the various portal clients that are attached to it.

The TEMA runs as a separate process from WebSphere Application Server instances.

The Data Collector runs within WebSphere Application Server instances.

Through the TEPS users can customize the monitoring process and perform certain actions on the monitored systems.

14.3.2 ITCAM for WebSphere application support installation

Before you can view data collected by monitoring agents, you must install and enable application support for those agents. Application support files provide

agent-specific information for workspaces, helps, situations, templates, and other data.

Configuring application support is a two-step process:

1. Installing the application support files (from installation media).

2. Enabling the application support (sometimes referred to as adding or activating the application support).

Installing application support on TEMS

You only need to install application support on TEMS once per monitoring agent type:

1. On the Tivoli Enterprise Monitoring Server (TEMS) server, from the folder where the code was extracted, launch the ITCAM for WebSphere installation script:

   `litstems:/nfshome/foulques/itcamwasagent # ./install.sh`

2. Interact with the script using the parameters shown in Table 14-1.

Table 14-1   Application support installation on TEMS parameters

Parameter	Value
IBM Tivoli Monitoring directory	/opt/IBM/ITM
Where to install	Install products to the local host
Product package to install	Tivoli Enterprise Monitoring Server support
Product to install	IBM Tivoli Composite Application Manager Agent for WebSphere Applications V07.10.00.00
Support package to re-seed	IBM Tivoli Composite Application Manager Agent for WebSphere Applications

Enabling (adding or activating) the application support on the TEMS is already done by the previous script when seeding.

Installing application support on TEPS

You only need to install application support on TEPS once per monitoring agent type:

1. On the Tivoli Enterprise Portal Server (TEPS) server, from the folder where the code was extracted, launch the ITCAM for WebSphere installation script:

   `litsteps:/nfshome/foulques/itcamwasagent # ./install.sh`

2. Interact with the script using the parameters in Table 14-2 on page 226.

Table 14-2   Application support installation on TEPS parameters

Parameter	Value
IBM Tivoli Monitoring directory	/opt/IBM/ITM
Where to install	Install products to the local host
Product package to install	Tivoli Enterprise Portal Browser Client support
Product to install	IBM Tivoli Composite Application Manager Agent for WebSphere Applications V07.10.00.00
Product package to install	Tivoli Enterprise Portal Server support
Product to install	IBM Tivoli Composite Application Manager Agent for WebSphere Applications V07.10.00.00

3. Enabling (adding or activating) the application support on the TEPS requires running the TEPS configuration utility completely. If the TEPS is already configured, run the configuration utility again, and press enter all the time to pick the existing default values:

```
litsteps:~ # /opt/IBM/ITM/bin/itmcmd config -A cq
```

4. Interact with the script using the parameters shown in Table 14-3.

Table 14-3   TEPS reconfiguration parameters

Parameter	Value
Edit the Common event console for IBM Tivoli Monitoring settings	Yes
ITM connector name	ITM1
Maximum number of events for this connector	100
View closed events	No
TEMS Host Name	litstems.ltic.pok.ibm.com
Which database do you want to use for TEPS	DB2
DB2 instance name	db2inst1
DB2 admin ID	db2inst1

Parameter	Value
TEPS DB2 database name	TEPS
TEPS DB2 database login ID	itmuser
Are you using DB2, Oracle or None for Warehouse?	NONE
Validate User with LDAP?	No

14.3.3 Interactively installing ITCAM for WebSphere

ITCAM for WebSphere is installed directly on the Linux guest where application servers are installed but not on the Deployment Manager Linux guest. To interactively install ITCAM for WebSphere:

1. Extract the code from the downloaded package or from the installation CD to the target Linux server where WebSphere Application Server is installed and configured:

   ```
   LITCRP01: /nfshome/foulques/itcamwasagent # tar -xvf 'ITCAM
   Application Diagnostics V7.1 ITCAM Agent for WebSphere Applications
   - CZAV9EN.tar'
   ```

2. On the WebSphere server, from the folder where the code was extracted, launch the ITCAM for WebSphere installation script:

   ```
   LITCRP01:/nfshome/foulques/itcamwasagent # ./install.sh
   ```

3. Interact with the script using parameters, as shown in Table 14-4.

Table 14-4   ITCAM for WebSphere installation parameters

Parameter	Value
IBM Tivoli Monitoring directory	/opt/IBM/ITM
Where to install	Install products to the local host
Product package to install	IBM Tivoli Monitoring components for this operating system
Product to install	IBM Tivoli Composite Application Manager Agent for WebSphere Applications V07.10.00.00

14.3.4 Interactively configuring the ITCAM WebSphere agent

To interactively configure the ITCAM WebSphere agent:

1. Launch the interactive configuration script for ITCAM for WebSphere:

   `LITCRP01:/opt/IBM/ITM/bin # ./itmcmd config -A yn`

2. Interact with the script using the parameters shown in Table 14-5.

Table 14-5 ITCAM WebSphere agent configuration parameters

Parameter	Value
Configuration type	Use this option to configure the Tivoli Enterprise Monitoring Agent (TEMA) port number or Agent ID. If you modify the Tivoli Monitoring Agent port, all Application Servers with Data Collectors must be restarted to complete the reconfiguration
Alternative Node ID	CRP01
Monitoring Agent will use this TCP socket port	63335
Will this agent connect to a TEMS?	Yes
TEMS Host Name	litstems.ltic.pok.ibm.com

3. Start the TEMA agent:

   `/opt/IBM/ITM/bin/itmcmd agent start yn`

14.3.5 Interactively configure the Data Collector

To interactively configure the Data Collector:

1. Launch the interactive configuration script for ITCAM for WebSphere:

   `LITCRP01:/opt/IBM/ITM/bin # ./itmcmd config -A yn`

2. Interact with the script using parameters, as shown in Table 14-6 on page 229.

Table 14-6   ITCAM WebSphere data collector configuration parameters

Parameter	Value
Configuration type	Use this option to configure the Data Collector to monitor application server instances. You can also use this option to configure the Data Collector to connect to the Managing Server. The option requires that either the Application Servers are running (WebSphere Application Server Base Edition) or the Node Agent and Deployment Manager are running (WebSphere Application Server ND or XD). The Servers must be restarted to complete the configuration.
Enable communication to Managing Server	False
Configuration mode	Default
WebSphere Type	WebSphere Application Server
WebSphere Profile Home	/opt/IBM/WebSphere/AppServer/profiles/Appsrv01
WebSphere Profile Name	AppSrv01
WebSphere Server Home	/opt/IBM/WebSphere/AppServer
Input Instance Name	cells/LITWSND1Cell01/nodes/LITCRP01Node01/servers/Cognos_Report_server
Input Server Instance Alias	CRP01
Input Administrative Server Host name	LITWSND1.ltic.pok.ibm.com
Input Server Administrative Port	8880
Do you want to use the user name and password stored in soap.client.props or sas.client.props of WebSphere?	Yes
TEMS Host Name	litstems.ltic.pok.ibm.com

3. Restart the WebSphere Application Server instance so that the Data Collector configuration is picked up by WebSphere.

14.3.6 Silently installing ITCAM for WebSphere

The silent installation can be used for automation and provisioning in the cloud. To silently install ITCAM for WebSphere:

1. Stop the existing running monitoring agent on the Linux instance.

2. Create a text file that contains the installation parameters for both the TEMA and the Data Collector.

```
/nfshome/foulques/itcamwasagent/silent_install_SAC.txt
INSTALL_ENCRYPTION_KEY=IBMTivoliMonitoringEncryptionKey
INSTALL_FOR_PLATFORM=ls3263
INSTALL_PRODUCT=yn
```

3. Launch the silent installation:

```
/nfshome/foulques/itcamwasagent # ./install.sh -q -h /opt/IBM/ITM -p
/nfshome/foulques/itcamwasagent/silent_install_SAC.txt
```

4. Restart the existing other monitoring agent that you might have stopped earlier on.

14.3.7 Silently configuring the agent and the Data Collector

The silent configuration can be used for automation and provisioning in the cloud. To silently configure the agent and the Data Collector:

1. Create a text file that contains all configuration parameters for both the TEMA and the Data Collector:

```
/nfshome/foulques/itcamwasagent/silent_config_SAC_CRP01.txt
CMSCONNECT=YES
HOSTNAME=litstems.ltic.pok.ibm.com
NETWORKPROTOCOL=ip.pipe
configure_type=tema_configure
KYN_ALT_NODEID=CRP01
KYN_PORT=63335
configure_type=dc_configure
J2EEMS_SELECT=False
ENABLE_TTAPI=no
was-type=was
KYN_WAS_HOME./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=/opt/IBM
/WebSphere/AppServer
KYN_WAS_SERVERS./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=cells
/LITWSND1Cell01/nodes/LITCRP01Node01/servers/Cognos_Report_server
KYN_APPSRVR_ALIAS./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=CRP
01
```

```
KYN_ADMIN_HOST./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=LITWSN
D1.ltic.pok.ibm.com
KYN_CONNECT_TYPE./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=SOAP
KYN_ADMIN_PORT./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=8879
KYN_USE_CLIENT_PROPS./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=
yes
KYN_USE_ALLOW_RECONFIG./opt/IBM/WebSphere/AppServer/profiles/Appsrv0
1=yes
KYN_GCLOG_PATH./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=${SERV
ER_LOG_ROOT}/itcam_gc.log
KYN_GCLOG_CYCLE./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=5,300
0
KYN_ITCAM_HS./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=enable
KYN_BACKUP_WAS./opt/IBM/WebSphere/AppServer/profiles/Appsrv01=false
```

2. Stop the TEMA agent:

   ```
   /opt/IBM/ITM/bin/itmcmd agent stop yn
   ```

3. Launch the silent configuration:

   ```
   /opt/IBM/ITM/bin/itmcmd config -A -p
   /nfshome/foulques/itcamwasagent/silent_config_SAC_CRP01.txt yn
   Agent configuration started...
   Agent configuration completed...
   ```

4. Start the TEMA agent:

   ```
   /opt/IBM/ITM/bin/itmcmd agent start yn
   Starting IBM Tivoli Composite Application Manager Agent for
   WebSphere Applications ...
   IBM Tivoli Composite Application Manager Agent for WebSphere
   Applications started
   ```

5. Restart WebSphere Application Server instance so that the Data Collector
   configuration is picked up by WebSphere.

14.4 Implementing the IBM HTTP Server agent

In this section, we describe the installation and the configuration of the
monitoring agent for the IBM HTTP Server.

14.4.1 Installing the ITCAM for HTTP application support

Before you can view data collected by monitoring agents, you must install and
enable application support for those agents. Application support files provide

agent-specific information for workspaces, helps, situations, templates, and other data.

Configuring application support is a two-step process:

1. Install the application support files (from installation media).

2. Enable the application support (sometimes referred to as adding or activating the application support).

Installing application support on TEMS

This only needs to be done once per monitoring agent type. To install application support on TEMS:

1. On the Tivoli Enterprise Monitoring Server (TEMS) server, from the folder where the code was extracted, launch the ITCAM for WebSphere installation script:

   ```
   litstems:/nfshome/foulques/itcamhttpagent # ./install.sh
   ```

2. Interact with the script using parameters, as shown on Table 14-7.

Table 14-7   Application support installation on TEMS parameters

Parameter	Value
IBM Tivoli Monitoring directory	/opt/IBM/ITM
Where to install	Install products to the local host
Product package to install	Tivoli Enterprise Monitoring Server support
Product to install	IBM Tivoli Composite Application Manager Agent for HTTP Servers V07.10.00.00
Support package to re-seed	IBM Tivoli Composite Application Manager Agent for HTTP Servers V07.10.00.00

Enabling (adding or activating) the application support on the TEMS is already done by the previous script when seeding.

Installing application support on TEPS

This only needs to be done once per monitoring agent type. To install the application support on TEPS:

1. On the Tivoli Enterprise Portal Server (TEPS) server, from the folder where the code was extracted, launch the ITCAM for HTTP installation script:

   ```
   litsteps:/nfshome/foulques/itcamhttpagent # ./install.sh
   ```

2. Interact with the script using parameters, as shown in Table 14-8.

Table 14-8   Application support installation on TEPS parameters

Parameter	Value
IBM Tivoli Monitoring directory	/opt/IBM/ITM
Where to install	Install products to the local host
Product package to install	Tivoli Enterprise Portal Browser Client support
Product to install	IBM Tivoli Composite Application Manager Agent for HTTP Servers V07.10.00.00
Product package to install	Tivoli Enterprise Portal Server support
Product to install	IBM Tivoli Composite Application Manager Agent for HTTP Servers V07.10.00.00

3. Enabling (adding or activating) the application support on the TEPS requires running the TEPS configuration utility completely. If the TEPS is already configured, run the configuration utility again and press Enter to pick the existing default values:

   ```
   litsteps:~ # /opt/IBM/ITM/bin/itmcmd config -A cq
   ```

4. Interact with the script using parameters, as shown in Table 14-9.

Table 14-9   TEPS reconfiguration parameters

Parameter	Value
Edit Common event console for IBM Tivoli Monitoring settings	Yes
ITM connector name	ITM1
Maximum number of events for this connector	100
View closed events	No

Parameter	Value
TEMS Host Name	litstems.ltic.pok.ibm.com
Which database do you want to use for TEPS	DB2
DB2 instance name	db2inst1
DB2 admin ID	db2inst1
TEPS DB2 database name	TEPS
TEPS DB2 database login ID	itmuser
Are you using DB2, Oracle or None for Warehouse?	NONE
Validate User with LDAP?	No

14.4.2 Interactively installing ITCAM for the HTTP agent

ITCAM for HTTP is installed directly on the Linux guest where HTTP servers are installed. To interactively install ITCAM for the HTTP agent:

1. Extract the code from the downloaded package or from the installation CD to the target Linux server where the HTTP Server is installed and configured:

   ```
   LITCIHS1: /nfshome/foulques/itcamhttpagent # tar -xvf 'ITCAM
   Application Diagnostics V7.1 ITCAM Agent for HTTP Servers -
   CZAW6EN.tar'
   ```

2. On the HTTP server, from the folder where the code was extracted, launch the ITCAM for HTTP installation script:

   ```
   LITCIHS1:/nfshome/foulques/itcamhttpagent # ./install.sh
   ```

3. Interact with the script using parameters, as shown in Table 14-10.

Table 14-10   ITCAM for HTTP installation parameters

Parameter	Value
IBM Tivoli Monitoring directory	/opt/IBM/ITM
Where to install	Install products to the local host
Product package to install	IBM Tivoli Monitoring components for this operating system
Product to install	IBM Tivoli Composite Application Manager Agent for HTTP Servers V07.10.00.00

14.4.3 Interactively configuring ITCAM for the HTTP agent

To interactively configure ITCAM for the HTTP agent:

1. Launch the interactive configuration script for ITCAM for WebSphere:

   `LITCIHS1:/opt/IBM/HTTPServer/conf # `**`/opt/IBM/ITM/bin/itmcmd config`**
   `-A ht`

2. Interact with the script using parameters, as shown in Table 14-6 on page 229.

Table 14-11   ITCAM for HTTP agent configuration parameters

Parameter	Value
Monitoring Method	Fixed
Fixed Interval between Collections (sec)	60
On Demand Maximum Sample Age (sec)	15
Ping the Web Sites	Disabled
Maximum Number of Agent's Log Events	100
Alias Name for Apache Web server	IHS1
Web server configuration file path	/opt/IBM/HTTPServer/conf/httpd.conf
Path to the executable	/opt/IBM/HTTPServer/bin/apachectl
Arguments for the executable	<nothing>
TEMS Host Name	litstems.ltic.pok.ibm.com

3. Start the ITCAM for HTTP agent:

   `/opt/IBM/ITM/bin/itmcmd agent start ht`

4. Restart the IBM HTTP Server instance so that the HTTP server configuration is picked up.

14.4.4 Silently installing ITCAM for HTTP

The silent installation can be used for automation and provisioning in the cloud. To silently install ITCAM for HTTP:

1. Stop existing running monitoring agents, if any, on the Linux instance.

2. Create a text file that contains the installation parameters for ITCAM for HTTP.

```
/nfshome/foulques/itcamhttpagent/silent_install_SAC.txt
INSTALL_ENCRYPTION_KEY=IBMTivoliMonitoringEncryptionKey
INSTALL_FOR_PLATFORM=ls3263
INSTALL_PRODUCT=ht
```

3. Launch the silent installation:

```
LITCIHS2:/nfshome/foulques/itcamhttpagent # ./install.sh -q -h
/opt/IBM/ITM -p
/nfshome/foulques/itcamhttpagent/silent_install_SAC.txt
```

4. Restart existing monitoring agents that you stopped earlier.

14.4.5 Silently configuring ITCAM for HTTP

The silent configuration can be used for automation and provisioning in the cloud. To silently configure ITCAM for HTTP:

1. Create a text file that contains all configuration parameters for ITCAM for HTTP:

```
/nfshome/foulques/itcamhttpagent/silent_config_SAC_IHS2.txt
CMSCONNECT=YES
HOSTNAME=litstems.ltic.pok.ibm.com
NETWORKPROTOCOL=ip.pipe
[KHT_AGENT_CONFIG]
KHT_APACHE_WEBSRVR_CONFIG_PATH.IHS2=/opt/IBM/HTTPServer/conf/httpd.c
onf
KHT_APACHE_WEBSRVR_ALIAS.IHS2=IHS2
KHT_APACHE_EXEC_PATH.IHS2=/opt/IBM/HTTPServer/bin/apachectl
```

2. Stop the ITCAM for HTTP agent:

```
/opt/IBM/ITM/bin/itmcmd agent stop ht
```

3. Launch the silent configuration:

```
/opt/IBM/ITM/bin/itmcmd config -A -p
/nfshome/foulques/itcamhttpagent/silent_config_SAC_IHS2.txt ht
Agent configuration started...
KHTC0010I Info: /opt/IBM/HTTPServer/conf/httpd.conf was configured.

KHT_AGENT_CONFIG]
Agent configuration completed...
```

4. Start the ITCAM for HTTP agent:

```
/opt/IBM/ITM/bin/itmcmd agent start ht
```

```
Starting IBM Tivoli Composite Application Manager Agent for HTTP
Servers ...
IBM Tivoli Composite Application Manager Agent for HTTP Servers
started
```

5. Restart the IBM HTTP Server instance so that the configuration is picked up.

14.5 Implementing the DB2 database agent

In this section, we describe the installation and the configuration of the monitoring agent for DB2 database.

14.5.1 Installing ITM for database application support

Before you can view data that is collected by monitoring agents, you must install and enable application support for those agents. Application support files provide agent-specific information for workspaces, helps, situations, templates, and other data.

Configuring application support is a two-step process:

1. Install the application support files (from installation media).

2. Enable the application support (sometimes referred to as adding or activating the application support).

Installing application support on TEMS

This only needs to be done once per monitoring agent type. To install application support on TEMS:

1. On the Tivoli Enterprise Monitoring Server (TEMS) server, from the folder where the code was extracted, launch the ITM for Database installation script:

   ```
   litstems:/nfshome/foulques/itmdatabase # ./install.sh
   ```

2. Interact with the script using parameters, as shown in Table 14-12.

Table 14-12   Application support installation on TEMS parameters

Parameter	Value
IBM Tivoli Monitoring directory	/opt/IBM/ITM
Where to install	Install products to the local host
Product package to install	Tivoli Enterprise Monitoring Server support
Product to install	Monitoring Agent for DB2 V06.20.00.00
Support package to re-seed	Monitoring Agent for DB2 V06.20.00.00

Enabling (adding or activating) the application support on the TEMS is already done by the previous script when seeding.

Installing application support on TEPS

This only needs to be done once per monitoring agent type. To install application support on TEPS:

1. On the Tivoli Enterprise Portal Server (TEPS) server, from the folder where the code was extracted, launch the ITM for Database installation script:

   ```
   litsteps:/nfshome/foulques/itmdatabase # ./install.sh
   ```

2. Interact with the script using the parameters shown in Table 14-8 on page 233.

Table 14-13   Application support installation on TEPS parameters

Parameter	Value
IBM Tivoli Monitoring directory	/opt/IBM/ITM
Where to install	Install products to the local host
Product package to install	Tivoli Enterprise Portal Browser Client support

Parameter	Value
Product to install	Monitoring Agent for DB2 V06.20.00.00
Product package to install	Tivoli Enterprise Portal Server support
Product to install	Monitoring Agent for DB2 V06.20.00.00

3. Enabling (adding or activating) the application support on the TEPS requires running the TEPS configuration utility completely. If the TEPS is already configured, run the configuration utility again, and press Enter to pick the existing default values:

```
litsteps:~ # /opt/IBM/ITM/bin/itmcmd config -A cq
```

4. Interact with the script using the parameters shown in Table 14-9 on page 233.

Table 14-14   TEPS reconfiguration parameters

Parameter	Value
Edit Common event console for IBM Tivoli Monitoring settings	Yes
ITM connector name	ITM1
Maximum number of events for this connector	100
View closed events	No
TEMS Host Name	litstems.ltic.pok.ibm.com
Which database would you like to use for TEPS	DB2
DB2 instance name	db2inst1
DB2 admin ID	db2inst1
TEPS DB2 database name	TEPS
TEPS DB2 database login ID	itmuser
Are you using DB2, Oracle or None for Warehouse?	NONE
Validate User with LDAP?	No

14.5.2 Interactively installing ITM for the database agent

ITM for database is installed directly on the Linux guest where database servers are installed:

1. Extract the code from the downloaded package or from the installation CD to the target Linux server where the database server is installed and configured:

```
LITCDB2:/nfshome/foulques/itmdatabase # tar -xvf 'IBM Tivoli
Monitoring for Database V6.2 Fix Pack 1 Base Multiplatform -
C1R3JIE.tar'
```

2. On the database server, from the folder where the code was extracted, launch the ITM for database installation script:

```
LITCDB2:/nfshome/foulques/itmdatabase # ./install.sh
```

3. Interact with the script using parameters, as shown in Table 14-10 on page 234.

Table 14-15   ITM for Database installation parameters

Parameter	Value
IBM Tivoli Monitoring directory	/opt/IBM/ITM
Where to install	Install products to the local host
Product package to install	IBM Tivoli Monitoring components for this operating system
Product to install	Monitoring Agent for DB2 V06.20.00.00

14.5.3 Interactively configuring ITM for the database agent

To interactively configure ITM for the database agent:

1. Launch the interactive configuration script for ITM for Database:

```
LITCDB2:/nfshome/foulques/itmdatabase # /opt/IBM/ITM/bin/itmcmd
config -A ud
```

2. Interact with the script using parameters, as shown in Table 14-16.

Table 14-16   ITM for Database agent configuration parameters

Parameter	Value
TEMS Host Name	litstems.ltic.pok.ibm.com

3. Start the ITM for Database agent:

```
LITCDB2:/nfshome/foulques/itmdatabase # /opt/IBM/ITM/bin/itmcmd
agent -o db2inst1 start ud
itmcmd agent      : Sourcing db2profile for user db2inst1.
Starting Monitoring Agent for DB2 ...
Monitoring Agent for DB2 started
```

14.5.4 Silently install ITM for the database agent

The silent installation can be used for automation and provisioning in the cloud. To silently install ITM for the database agent:

1. Stop the existing running monitoring agent, if any, on the Linux instance.

2. Create a text file that contains the installation parameters for ITM for database:

```
/nfshome/foulques/itmdatabase/silent_install_SAC.txt
INSTALL_ENCRYPTION_KEY=IBMTivoliMonitoringEncryptionKey
INSTALL_FOR_PLATFORM=ls3263
INSTALL_PRODUCT=ud
INSTALL_PRODUCT=or
INSTALL_PRODUCT=oy
INSTALL_PRODUCT=oq
```

3. Launch the silent installation:

```
/nfshome/foulques/itmdatabase # ./install.sh -q -h /opt/IBM/ITM -p
/nfshome/foulques/itmdatabase/silent_install_SAC.txt
```

4. Restart existing monitoring agents.

14.5.5 Silently configuring ITM for database

The silent configuration can be used for automation and provisioning in the cloud. To silently configure ITM for database:

1. Create a text file that contains all configuration parameters for ITM for Database:

```
/nfshome/foulques/itmdatabase/silent_config_SAC_DB2.txt
CMSCONNECT=YES
HOSTNAME=litstems.ltic.pok.ibm.com
NETWORKPROTOCOL=ip.pipe
```

2. Launch the silent configuration:

```
/opt/IBM/ITM/bin/itmcmd config -A -p
/nfshome/foulques/itmdatabase/silent_config_SAC_DB2.txt ud
```

```
Agent configuration started...
Agent configuration completed...
```

3. Start the ITM for Database agent:

```
/nfshome/foulques/itmdatabase # /opt/IBM/ITM/bin/itmcmd agent -o
db2inst1 start ud
itmcmd agent       : Sourcing db2profile for user db2inst1.
Starting Monitoring Agent for DB2 ...
Monitoring Agent for DB2 started
```

14.6 Implementing the operating system monitoring agent

In this section, we describe the installation and the configuration of the monitoring agent for the Linux operating system.

IBM Tivoli Monitoring for Operating System agent is installed directly on the Linux for System z virtual server where the SLES11 is running.

Before the installation starts, extract the code from the downloaded package to the target Linux server. See Example 14-1 for a sample tar command to extract the ITM install package.

Example 14-1   Sample tar command to extract the ITM O/S agent install package

```
tar -xvf ITMv622_OS_Agent_CZ8XZEN.tar
```

The subsequent sections described the steps that are used to install and configure a distributed monitoring agent on our lab Linux on System z virtual servers.

14.6.1 Interactively installing the operating system monitoring agent

In the directory where you extracted the installation files, run the `install.sh` command. The installation program automatically detects the operating system type version and level. Table 14-17 shows the values, selection, and options that we chose during the interactive installation.

Table 14-17   Values used for the installation script

Selection / option	Value used
Target install directory	/opt/IBM/ITM

Selection / option	Value used
Install option	Install products to the local host
Encryption key	IBMTivoliMonitoringEncryptionKey
Agent install	Monitoring Agent for Linux OS

14.6.2 Configuring the operating system monitoring agent

To configure the operating system monitoring agent:

1. After the installation completes, run the following two commands to configure the agent:

   – SetPerm: Changes the ownership of additional agent files. Table 14-18 shows the value that we chose for the SetPerm command.

Table 14-18   Value used for the SetPerm command

Install option	Value used
Product	Monitoring Agent for Linux OS Linux S390 R2.6

   – itmcmd config: Configures your monitoring agent. Table 14-19 shows the values we chose for the agent configuration.

Table 14-19   Values used for the agent configuration

Configuration option	Value used
connect to a TEMS	Yes
network protocol	ip.pipe
network protocol 2	0
IP.PIPE port number	1918
KDC_PARTITION	null
connection for a secondary TEMS	No
Optional Primary Network Name	0

14.6.3 Verifying installation

The `cinfo` command can be used to verify that the product was installed successfully, as shown in Example 14-2 on page 244.

Example 14-2   Sample ITM O/S agent product information

```
LITCBA02:/opt/IBM/ITM/bin # ./cinfo -i

*********** Mon Apr 19 15:15:51 EDT 2010 ******************
User: root Groups: root
Host name : LITCBA02     Installer Lvl:06.22.01.00
CandleHome: /opt/IBM/ITM
********************************************************
...Product inventory

ax      IBM Tivoli Monitoring Shared Libraries
          ls3266  Version: 06.22.01.00

gs      IBM GSKit Security Interface
          ls3263  Version: 07.40.20.00
          ls3266  Version: 07.40.20.00

jr      Tivoli Enterprise-supplied JRE
          ls3266  Version: 05.09.00.00

lz      Monitoring Agent for Linux OS
          ls3266  Version: 06.22.01.00

ui      Tivoli Enterprise Services User Interface
          ls3266  Version: 06.22.01.00
```

14.6.4 Silently installing the operating system monitoring agent

To run a silent install of the Operating System monitoring agent:

1. Update the response file silent_install.txt to find the directory where you extracted the installation package. Example 14-3 shows the response file name and its location.

Example 14-3   Sample response file name and location

```
/nfshome/simon/ITMv622Agent/silent_install.txt
```

2. Update the parameters in the response file, as shown in Example 14-4.

Example 14-4   Sample response file content

```
INSTALL_PRODUCT=lz
```

3. Run the silent **install** command, as shown in Example 14-5 on page 245.

> **Note:** Enter the `install` command line all in one line.

Example 14-5   Sample Silent install command

```
./install.sh -q -h /opt/IBM/ITM -p
/nfshome/simon/ITMv622Agent/silent_install.txt
```

4. After the install completes successfully, update the response file to configure the operating system monitoring agent in silent mode, as shown in Example 14-6.

Example 14-6   Sample response file to configure the O/S monitoring agent

```
CMSCONNECT=YES
HOSTNAME=LITSTEMS
NETWORKPROTOCOL=ip.pipe
IPPIPEPORTNUMBER=1918
PORTNUMBER=1918
```

5. Run the `CandleConfig` command in silent mode, as shown in Example 14-7.

> **Note:** Enter the `config` command line all in one line.

Example 14-7   Sample CandleConfig command

```
cd /opt/IBM/ITM/bin
./CandleConfig -A -p /nfshome/simon/ITMv622Agent/silent_config.txt lz
```

14.6.5 Operating System monitoring agent startup and shutdown

After the agent is configured. The agent can be started. It will connect to the TEMS. Example 14-8 shows the agent startup and a simple verification for the network connection between the TEMS and the O/S agent.

Example 14-8   Sample ITM O/S agent start up and quick verification

```
LITCBA02:/opt/IBM/ITM/bin # ./itmcmd agent start lz
Starting Monitoring Agent for Linux OS ...
Monitoring Agent for Linux OS started

LITCBA02:/opt/IBM/ITM/bin #
LITCBA02:/opt/IBM/ITM/bin #ps axf|grep ITM
17982 pts/0    Sl     0:00 /opt/IBM/ITM/ls3266/lz/bin/klzagent
19530 pts/0    Sl     0:00 /opt/IBM/ITM/ls3266/lz/bin/kcawd
```

```
LITCBA02:/opt/IBM/ITM/bin # netstat -an|grep 192.168.71.109
tcp 0 0 192.168.71.89:45844     192.168.71.109:1918     ESTABLISHED
```

To stop the agent, issue the command, as shown in Example 14-9.

Example 14-9   Sample command to stop the O/S monitoring agent

```
cd /opt/IBM/ITM/bin
./itmcmd agent stop lz
```

Part 5

Driving the cloud

Cloud computing is an emerging style of IT delivery in which applications, data, and IT resources are rapidly provisioned and provided as standardized offerings to users. It is also a way of managing large numbers of highly virtualized resources such that, from a management perspective, they resemble a single, large resource. Therefore providing a cloud includes not only hardware and software but also Service Management. In this part, we discuss the main aspects that must be considered when providing and managing a cloud, such as monitoring, automation, provisioning, requesting a new cloud service and onboarding, capacity management, and security implemented as a Smart Analytics Cloud in our lab environment.

Service life cycle

Service Management is another key area of the cloud. When providing a cloud service, the entire life-cycle from service definition through service termination must be managed and supported by tools where ever possible. Best practices enhance the quality of the service provided. Figure 15-1 on page 250 shows the stages of the Service Life Cycle for a cloud service.

The service life-cycle includes:

► Definition: In this phase, the services that are provided with the Smart Analytics Cloud are defined. The required components are selected here and the initial design is completed. This phase follows the architectural process as it is described in Part 3, "Architecture" on page 45.

► Offering: After the service is defined, it is published in the service catalog and end users can select and order them.

► Onboarding and provisioning: After the request is initiated, the automatic process or workflow for setting up the components to provide the Smart Analytics Cloud service starts:

  – Infrastructure: Set up the systems

  – Software: Install the required Linux guests on z/VM, WebSphere Application Server, DB2, and so on

  – Applications, such as, Cognos 8 BI and Monitoring

– Configuration: Set up IP configurations, user IDs, roles, and rights. After implementation, the application can be used by the end-users.

► Production: This includes the maintenance, support, and operations of the solution. Monitoring is also done in this phase.

► Termination: Deprovision resources when users terminate the services. This phase is not so important when setting up a productive Smart Analytics Cloud, but when offering other types of clouds, such as a test cloud, which is terminated after the test phase has completed this phase, must be defined too.
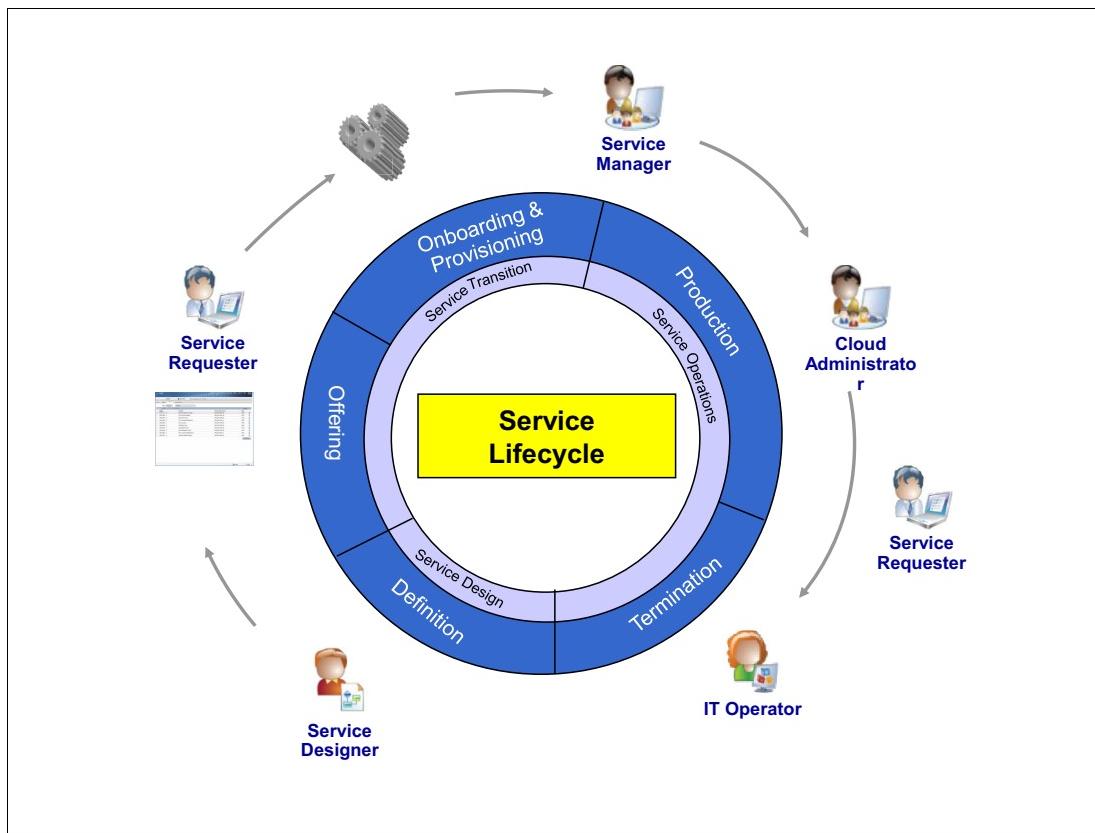
Figure 15-1   Service life-cycle for cloud

When we think of managing the cloud environment, managing really comes down to three focal points: visibility, control, and automation:

► Visibility: The ability to see everything that is going on in the cloud infrastructure. Users must have an integrated, actionable, and insightful view into critical metrics.

► Control: The ability to keep the cloud infrastructure in its desired state by enforcing policies. It means to improve process discipline and remain effective as systems grow.

► Automation: The ability to manage huge and growing cloud infrastructures and controlling cost and quality. It includes the improvement of quality and reducing costs through operational and workflow automation.

Because the cloud is a service delivery model, the service management solution is necessary to provide cloud computing efficiently.

The system management solution used in our lab environment is based on these focal points. We used IBM Tivoli components to implement the concept. Although Tivoli's entire Service Management Portfolio is applicable to Cloud Management, we focus on three main areas in this IBM RedBooks publication:

► Monitoring the cloud infrastructure: Discovers and monitors the virtual resources.

► Provisioning the cloud infrastructure: Enables an automated deployment of the virtual resources, such as machines, storage, and networks, but also the application components, such as IBM WebSphere Application Server and IBM Cognos 8 BI.

► Managing the cost of the cloud infrastructure: Determines the cost of service delivery.

16

Onboarding

To provide a cloud effectively and efficiently a service process management platform is required. It must deliver and automate the dynamic infrastructure and provide self service for the end user.

An onboarding application allows end users to use IT services without being an expert in IT. It improves customer satisfaction by accelerating service delivery with automated work flows.

The onboarding application that is used in our lab environment was described earlier. This chapter exceeds the functionality of the lab application and provides a perspective on the complete functionality that an onboarding application must provide.

The quality of service delivery is improved because standardized delivery models and templates are used. It allows faster and consistent service deployment because the service offerings are directly available to the consumers. Integrated and role-based process work flows are used. Automation reduces the time to deploy, IT operational costs, and errors.

Situation without using automated workflow and onboarding process:

► Complete paper request

► Call IT daily to check status

► Hope hardware is available

253

- ► Provisioning is manual and inconsistent
- ► When system resources are needed only an out-of-date development environment is available
- ► During production deployment, IT discovers that the application will not run in the current production OS
- ► Meet project demands. The out-of-date environment configuration is put in production to satisfy the critical business need

Situation with automated workflow and onboarding process:

- ► Automated with self service portal
- ► Track workflow status online
- ► Services when you need them
- ► Provisioning is automated with implemented standards
- ► The service catalog contains standardized images and environments that are automatically updated
- ► The developers get a production standard environment when needed with an outstanding user experience using the self serve portal

Figure 16-1 on page 256 illustrates the onboarding process.

16.1 Roles in the onboarding process

The roles involved in the onboarding process are:

- ► Service Requester
- ► Service Request Approver
- ► Service Manager
- ► Smart Analytics Cloud Administrator
- ► Smart Analytics Cloud Consultant

The *Service Requester* can be any person in the large enterprise who can use the services of the Smart Analytics Cloud. The Service Requester has access to the Service Catalog and the onboarding application and can order an item from the Service Catalog.

The *Service Request Approver* has the permission to approve the service of the requested by the Service Requester. The *Approver* is a role that has financial and business scope approval rights within the enterprise. Depending on the scope of the requested service, appropriate approval levels must be assigned to that role.

The *Service Manager* is the interface between the demand side (Service Requester and Service Request Approver) and the supply side (Smart Analytics Cloud Administrator and Smart Analytics Cloud Consultant). This role coordinates and plans the activities on the supply side that are necessary to provide the requested service and makes sure that the service is provided in the required quality.

The *Smart Analytics Cloud Administrator* is responsible for planning and providing the technical implementation of the Smart Analytics Cloud.

When service request includes other activities besides just implementing a technical environment, such as, process definitions, report design, the *Smart Analytics Cloud Consultant* performs these activities.

16.2 Onboarding process

The onboarding process consists of a planning and execution phase.

Planing phase
Figure 16-1 on page 256 illustrates the planning phase.
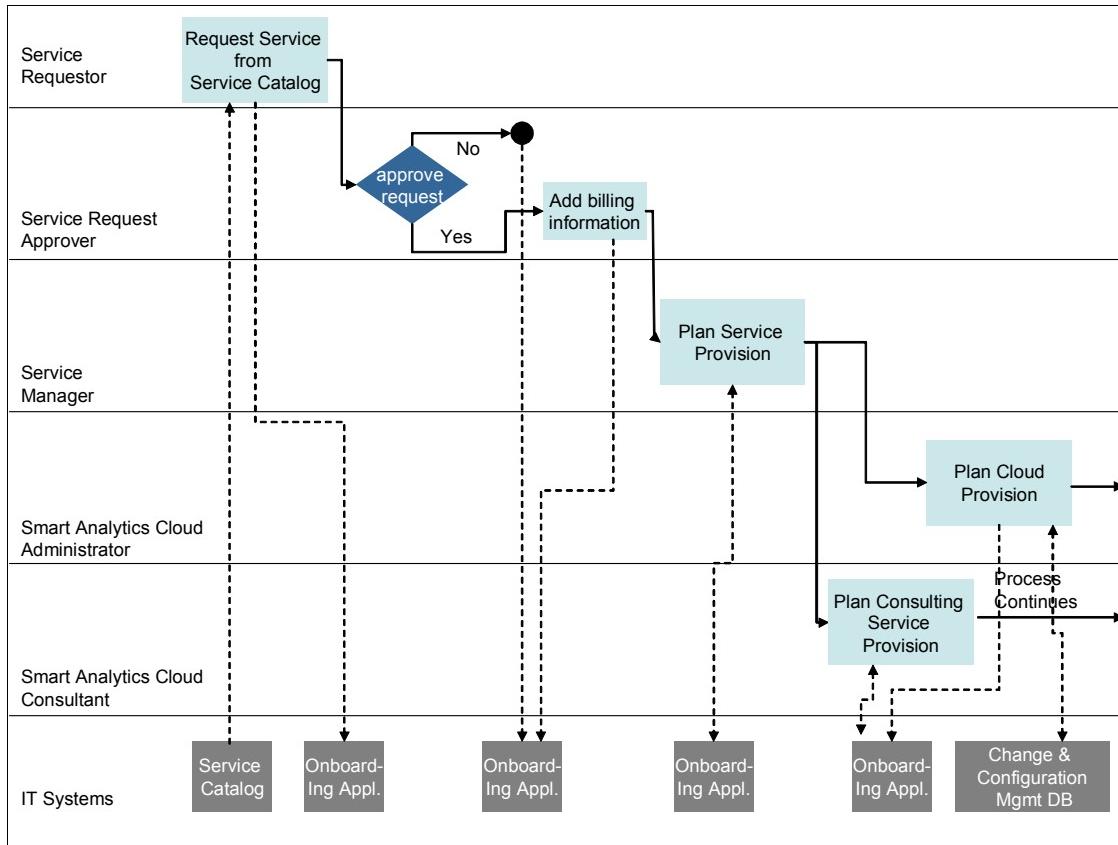
Figure 16-1   Onboarding process: Planning

The onboarding process starts when the *Service Requester* requests a Service from the Service Catalog, which is a repository that contains all Services provided by the IT organization. It can contain hardware, software, implementation, and consulting services or any combination of them. The services provided by a Smart Analytics Cloud, as detailed in Chapter 4, "Scope of the Smart Analytics Cloud" on page 27, are also available in the service catalog.

The *Service Requester* can be any person in the enterprise that can potentially have access to the services of the Smart Analytics cloud. The request is initiated in the onboarding application that is described in 6.2, "Onboarding application" on page 66.

After the request is initiated and all required fields in the onboarding application are populated, the *Service Request Approver* evaluates the service request and approves or rejects it. The Service Request Approver is a management role that

judges over the business needs of the Service Request and is responsible for the financial aspects.

When the Service Request is rejected, the decision is documented in the onboarding application and the onboarding process ends.

When the Service Request is approved, the billing information is added and documented in the onboarding application and the *Service Manager* proceeds.

The *Service Manager* starts to plan the supply of the service. For the Smart Analytics Cloud, it must be checked when the *Smart Analytics Cloud Administrator* has time available to complete the request and whether further purchases (additional licences, storage, CPU) must be made. When the requested Service includes consulting services, these also must be planned and available.

When the overall planning is completed the S*mart Analytics Cloud Administrator* and the *Smart Analytics Cloud Consultant* plan their tasks.

The activity, Plan Cloud Provision, the *Smart Analytics Cloud Administrator* plans the capacity and collects the necessary security and monitoring information using the Configuration Management Database.

The activity, Plan Consulting Service Provision, includes defining the scope and deliverables for the consulting activities and creating a project plan.

Execution phase
Figure 16-2 on page 258 illustrates the execution phase.
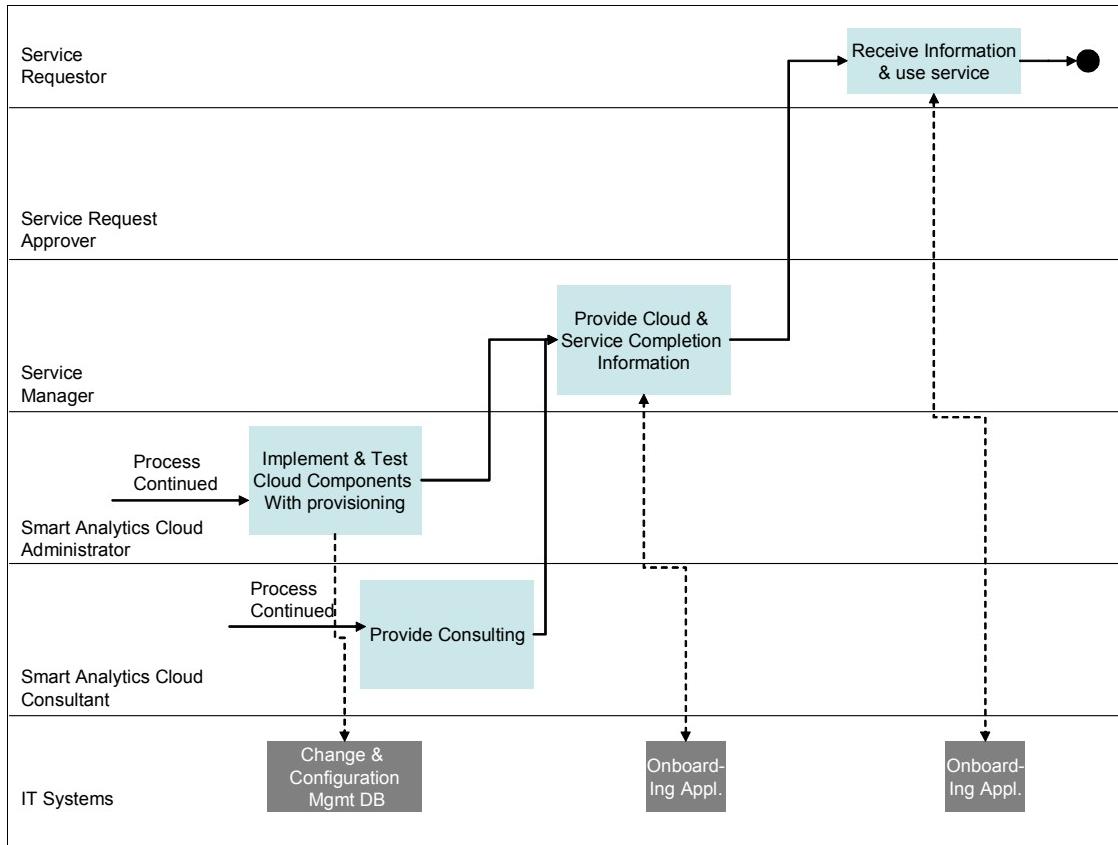
Figure 16-2   Onboarding process: Execution

After the planning is completed, the *Smart Analytics Cloud Administrator* implements the required hardware and software, including systems management software. Automatic provisioning is used where ever possible.

The *Smart Analytics Cloud Consultant* proceeds with the planned activities.

When the consulting activities are complete and the Smart Analytics components are successfully implemented, tested, and documented in the Configuration Management Database and the onboarding application, the *Service Manager* notifies the *Service Requester* of the result.

The *Service Requester* receives the information and can now use the provided Service. The onboarding process ends at this point.

The process described does not cover the necessary steps for planning and setting up the Smart Analytics Cloud. However, the most important aspects are

described in Chapter 17, "Provisioning and resource management" on page 261, Chapter 18, "Monitoring" on page 269, and Chapter 19, "Metering and billing" on page 279.

17

Provisioning and resource management

Efficient cloud computing depends heavily on fast provisioning. In our lab environment, we chose the IBM Workload Mobility Workbench, which provides an intuitive Internet interface for automatic provisioning of operating systems and software. In this chapter, we demonstrate an example of provisioning a new server with DB2 Database version 9.5 and discuss the necessary steps for on-the-fly resource management.

17.1 Provisioning a new server with DB2 Database

To create a new Linux guest on System z with a pre-installed DB2 Database we must complete eight steps, which are shown in Figure 17-1. The first five of them design the request and the last three describe where the infrastructure is placed.

The special icons placed at the left side of each menu's item inform about completion status.

Figure 17-1   Typical menu of new server provisioning

The *Administration* section contains details about the process surrounding the build request. A set of information is completed that stores the information about who initiated requests for provisioning, for what purpose, and so on. This does not have a direct impact on the technical side of provisioning.

Figure 17-2 on page 263 shows the available choices of the operating system and middleware. Select **s390x** in the architecture field for the System z platform. It is well known in the Linux world as a synonym for z/Architecture® (you can see this output by using `uname -p` in the Linux command line window).

The list of available options is based on an initial selection of architecture and operating system. Changing these parameters resets all software selections.

The list of available options is a representation of the Provisioning Assets Library, so add the middleware there first before you start the provisioning process. More information is available in 13.3, "Preparing software images" on page 215.

Figure 17-2   Choosing version of OS and middleware for provisioning

The initial list of Local File systems, which you can see in Figure 17-3 on page 264, includes defaults that are associated with the selected operating system and middleware. You can add additional file systems, which you want to be attached at the moment of provisioning. Using the DirMaint facility of z/VM you do not need to think about space management because it will attach DASD devices from predefined ranges that it maintains automatically.

Figure 17-3   Configuring local file systems for a new server

The Local User IDs step defines users and groups, which are created with a new system. You can manually specify typical parameters for Linux users, such as shell and home directory. The same as with file systems, user lists, by default, include users that are defined with software packages, in our case db2inst1 and db2fenc1.

In the Comments step, provide additional descriptions of the new system. This information is useful in large environments, especially when several people operate the data center.

On the Target Hardware page, shown in Figure 17-4 on page 265, you can see the configuration of the target virtual machine. Select the necessary amount of memory and virtual CPUs.

Note: Do not assign too much memory for Linux guest. Make it small because it works well without swapping. Avoid defining more virtual CPUs than physical ones that exist in the system. You can find other performance hints and tips at:

http://www.ibm.com/developerworks/linux/linux390/perf/index.html

Figure 17-4   Specifying target hardware configuration for a new server

On the Target Host panel, choose **z/VM** on which new virtual machine will be deployed. In the next step, you can change the Target Network configuration and specify the IP address and host name.

After you complete all eight steps, a Summary page is displayed, as shown in Figure 17-5 on page 266. It gives an overview of all your decisions. Make sure that you completed everything correctly.

Figure 17-5   Summary of provisioning request for new server

17.2 Reallocating resources on-the-fly

Reallocating resources between guest machines or adding them after a System z upgrade, in most cases, can be done on-the-fly without interrupting the system. Some preparations must be done before this feature can be used, so refer to 13.4, "Dynamic resource reallocation planning" on page 218 for information about.

17.2.1 Reallocating CPU resources

To dynamically add one or more logical processors to the logical partition:

1. Open the Hardware Management Console.

2. Select a CPC image.

3. Open the **Logical Processor Add** task for an active partition.

4. Change the logical processor definitions for this partition.

5. To have the new changes take effect immediately, click **Change Running System**.

To do the following for z/VM guest virtual machine, you must use the `set share` command to activate the stopped CPU.

17.2.2 Managing dynamic memory

Use the `query storage` command to check how much initial memory is defined in an LPAR profile that is reserved and in use now. To increase usage of memory by z/VM, use the `set storage` command. Refer to *CP Planning and Administration*, SC-24-6083-07 for more information.

17.2.3 Managing dynamic Input/Output

To make Input/Output (I/O) changes on the partition level you must have the Hardware Configuration Definition (HCD) service installed and configured on z/VM or z/OS in any of the LPARs. Using the HCD you can describe new LCU or OSA cards and give or reduce an access to them for one of LPARs and many other options.

18

Monitoring

In this chapter, we describe the monitoring of the Smart Analytics Cloud that we implemented in our lab environment. We show various panels that can be used to understand the current health of the cloud and be proactive for future needs. We also show how the event management with situations can facilitate pinpointing issues for faster resolution.

18.1 Monitoring cloud availability and health

Similar to the traditional IT environment, within the cloud environment, there is also a need to understand the performance trends for virtual and physical servers to determine capacity limitations and balance workloads. Alerts are required when problems with virtual resources occur that must be isolated and fixed quickly.

The following common areas require monitoring:

► Physical and virtual operating system resources, such as CPU, I/O, network, and LPAR

► Web resources, including web servers and application servers

► Middleware applications, such as Cognos

► Databases, such as DB2

► Transaction performance end-to-end across multiple virtual systems to isolate bottlenecks quickly

We use the following Tivoli products in our environment for monitoring:

► Tivoli Enterprise Monitoring Server (TEMS)
► Tivoli Enterprise Portal Server (TEPS)
► IBM Tivoli Monitoring (ITM) agent for Operating System
► IBM Tivoli Monitoring (ITM) agent for Database
► ITCAM Application Diagnostics agent for HTTP Servers
► ITCAM Application Diagnostics agent for WebSphere Applications

18.1.1 z/VM monitoring

The metrics about z/VM are provided to Tivoli Monitoring by Tivoli OMEGAMON XE on z/VM and Linux.

It provides a wide range of information about the z/VM operating system, its resources, and workloads. Information about Linux instances running as z/VM guests and the Linux workloads reveal how these instances and workloads on Linux are performing and impacting z/VM and each other.

IBM Tivoli OMEGAMON XE on z/VM and Linux utilizes the data collection from the Performance Toolkit for VM (PTK is a prerequisite) and complements it with data collection by the ITM Linux for zSeries agent. The Performance Tool Kit is the foundation for gathering z/VM metrics. The Linux agent has been the basis for monitoring Linux on all platforms. OMEGAMON XE on z/VM and Linux takes

advantage of the Tivoli Enterprise Portal (TEP) and allows for all of the Tivoli Enterprise Portal alerting, action, and integration capabilities to be utilized.

A workspace is the working area of the Tivoli Enterprise Portal window. You can start monitoring activity and system status immediately with the predefined workspaces. With just a few clicks of the mouse, you can customize your own workspaces to give summary overviews or to look at specific conditions.

Figure 18-1 shows a sample view of the z/VM system workspace.
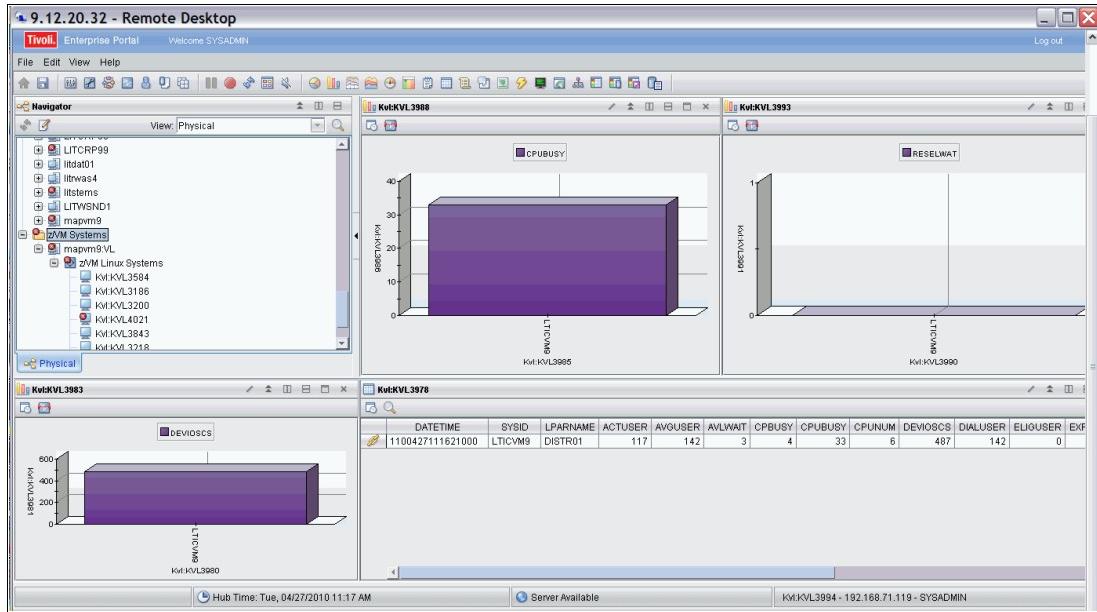
Figure 18-1   Tivoli Monitoring: z/VM system workspace

The product displays statistics for z/VM and its guest systems, including: System utilization, DASD, LPAR utilization, PAGING and SPOOLING utilization, REAL STORAGE utilization, TCP/IP utilization, Workload activity, CPU usage, I/O statistics, File system details, Disk and memory space, Network performance, and process information.

18.1.2 Linux monitoring

You can access the Linux Workload workspace to examine the performance of the Linux guests. The Linux Workload workspace provides data about system usage by user ID for all Linux guest systems that are defined in the cloud. The predefined workspace contains the Top 5 CPU Linux Guest Systems bar chart that displays the top five Linux guest users of CPU. For each Linux guest, it

shows the percentage of CPU used by the system to manage the workload, and the percentage of virtual CPU utilized. Figure 18-2 shows the Linux on System z workspace.
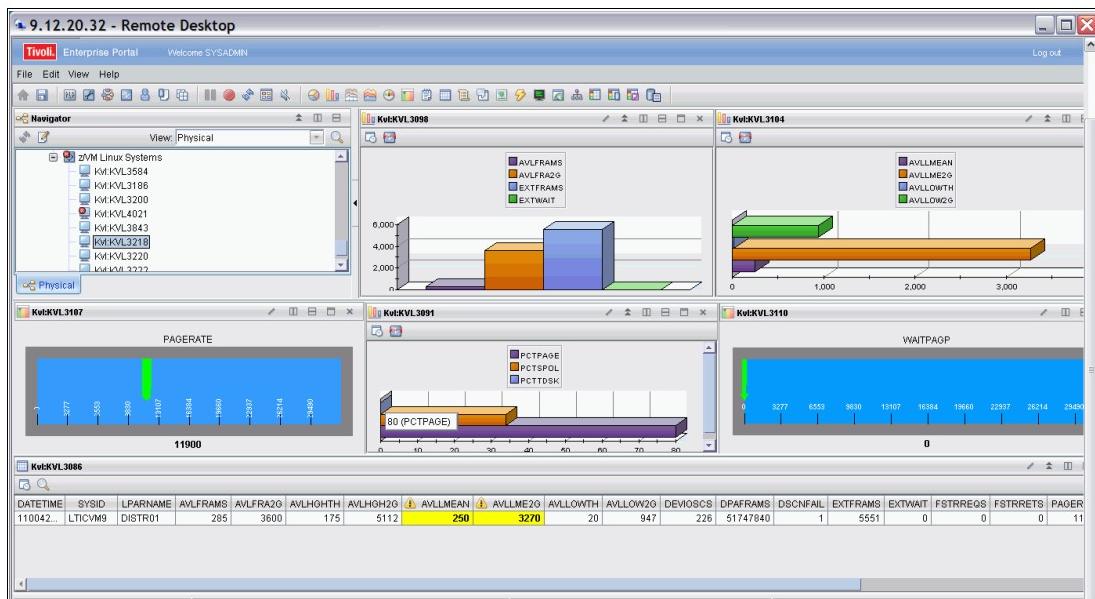
Figure 18-2   Tivoli Monitoring: Linux on System z workspace

Figure 18-2 shows a sample Linux for System z workspace.

The ApplData workspace, also accessed from the Workload workspace, provides extensive metrics about network utilization, CPU usage, and operating system data for each Linux guest system.

Figure 18-3 on page 273 shows detailed information about the Linux guest health.
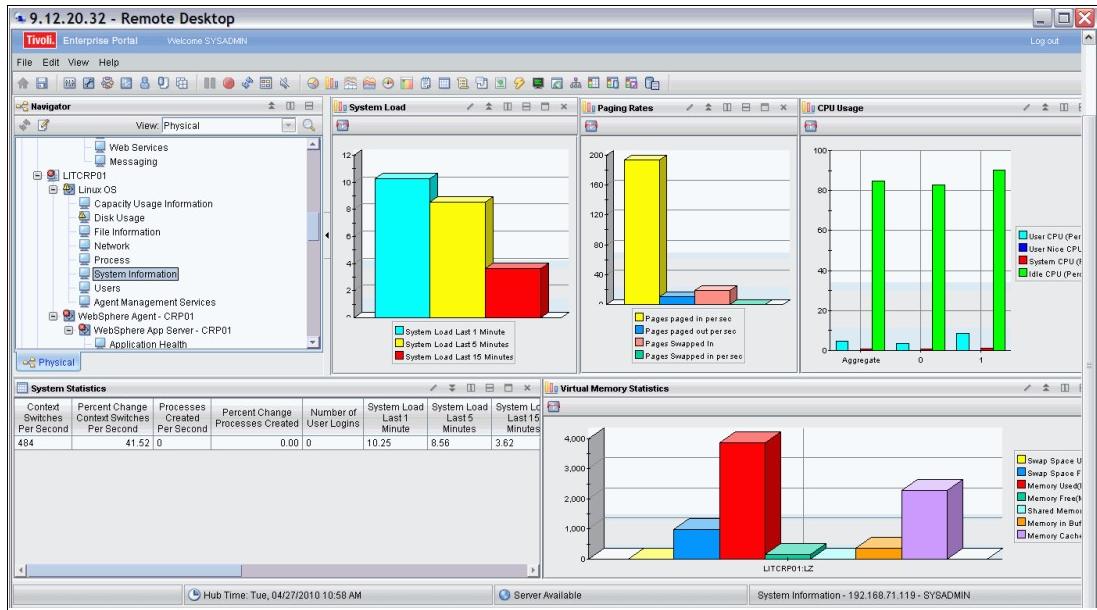
Figure 18-3   Tivoli Monitoring: Linux operating system workspace

18.1.3 WebSphere Application server monitoring

Most Cognos application components run on top of WebSphere Application Server. It is important to monitor the availability and the health of WebSphere.

Workspaces offer views of monitoring data that provide detailed current information about the WebSphere application servers that are running in the cloud.

Many predefined workspaces are provided with the ITCAM for WebSphere agent and allows you to monitor all aspects of the WebSphere server's health.

There are four types of data collected by ITCAM for the WebSphere agent:

► Resource Data: Obtained from Performance Monitoring Infrastructure (PMI) modules provided by WebSphere Application Server. They consist of aggregated data about how resources in the system are being used. They populate the following TEP workspaces: Pool Analysis, Web Applications, EJB Containers, DB Connection Pools, J2C connection Pool, Thread Pools, Cache Analysis, Workload Management, Scheduler, Web Services, and Platform Messaging.

► Request Data: Obtained from the Byte Code Injection (BCI) of WebSphere and application classes by the ITCAM Data Collector. They mainly track

individual user transactions and the elapsed time to perform various operations. They populate the following TEP workspaces: Application Health, Request Analysis, data sources, and JMS Summary.

► Garbage Collection Data: Obtained from parsing verbose GC file (on WebSphere Application Server/DC server). They are displayed in the Garbage Collection Analysis workspace.

► WebSphere Log Entries: Obtained from parsing WebSphere Application Server logs (on TEMA server). They are displayed in the Log Analysis workspace.

Data is always *pulled* by the agent from the Data Collector and can be obtained either at Fixed Interval or On-demand configuration. Figure 18-4 shows one of the many WebSphere workspace.
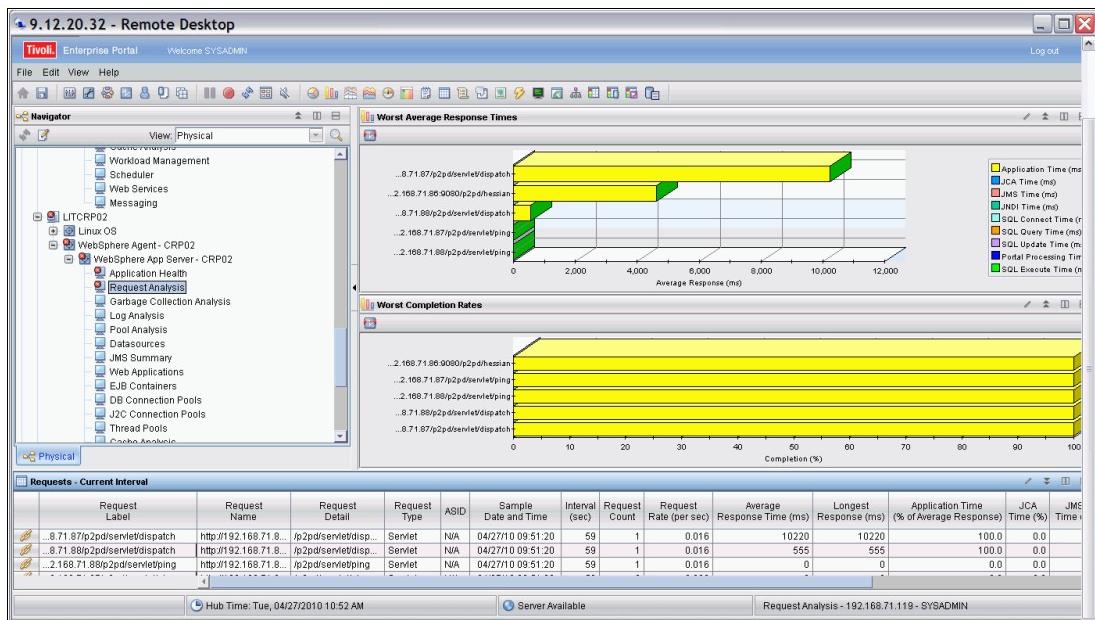
Figure 18-4   Tivoli Monitoring: WebSphere Application Server workspace

18.1.4 HTTP server monitoring

All incoming HTTP requests to Cognos go through the IBM HTTP servers first. The HTTP servers running the WebSphere plug-in do the load-balancing of requests to the Cognos application servers. These are good entry points to monitor to evaluate the HTTP incoming traffic to the Smart Analytics Cloud.

With the TEP workspaces for the ITCAM Agent for HTTP Servers, you can see into the data that is reported for HTTP Servers. The workspaces provide status, definitions, and statistical information about the HTTP Servers.

The Apache website's workspace shows the lists of the web sites (virtual hosts) configured for the Apache Web server and IBM HTTP Server along with their status and activity information, as shown in Figure 18-5.
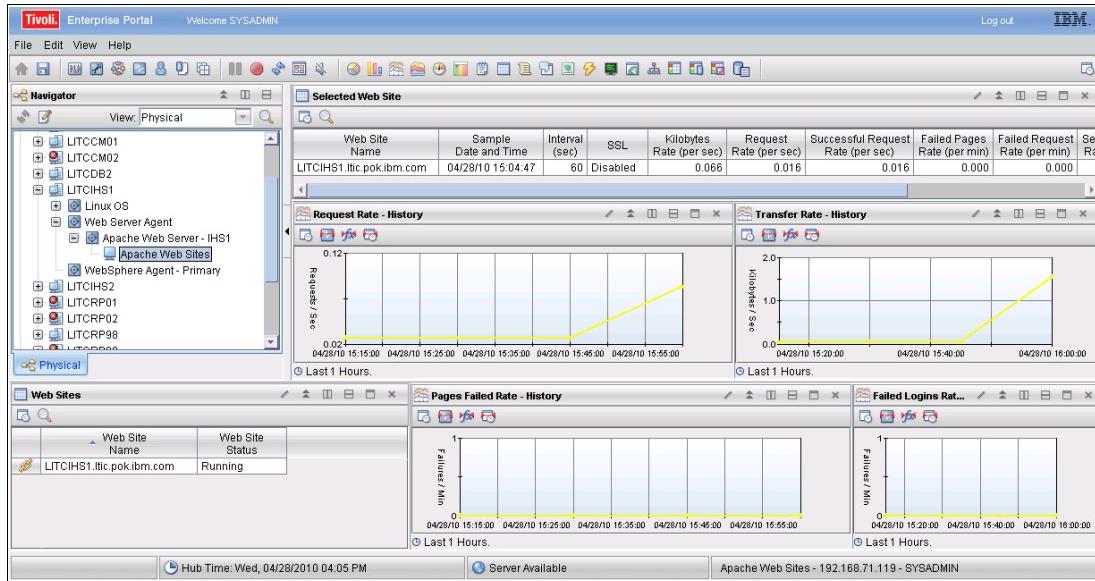
Figure 18-5   Tivoli Monitoring: IBM HTTP Server workspace

18.1.5 DB2 Database monitoring

Using the monitoring agent for DB2 you can easily collect and analyze specific information, including information about:

► Applications with the highest percentage of failed SQL statements, sort overflows, lock timeouts and deadlocks, and the lowest buffer pool hit ratio

► Buffer pool hit ratio by buffer pool, buffer pool hit ratio by database, average read and write times, asynchronous and synchronous I/O activity, extended store and non-buffer pool I/O activity

► Databases with the highest percentage of failed SQL statements, the lowest buffer pool hit ratio, and the highest number of connections, lock timeouts, and deadlocks

► Applications currently waiting for locks and other details about lock resources

► Server key events, the number of server connections, the databases with the lowest buffer pool hit ratio, and applications with the highest percentage of failed SQL statements

► Table spaces

Figure 18-6 shows one of the many DB2 workspace.
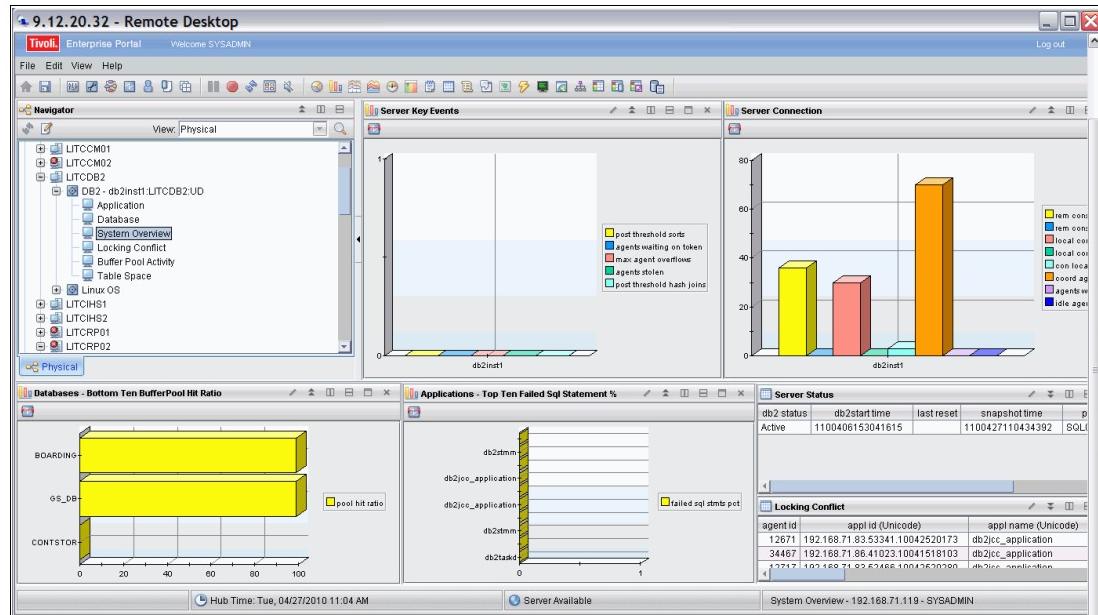
Figure 18-6   Tivoli Monitoring: DB2 system overview workspace

18.1.6 Cognos monitoring

Cognos is an application that runs within WebSphere Application Server. Consequently certain workspaces from the ITCAM for WebSphere agent relate to the Cognos application specifically and allows you to check the health of Cognos.

Figure 18-7 on page 277 shows a workspace that applies to the Cognos application specifically.
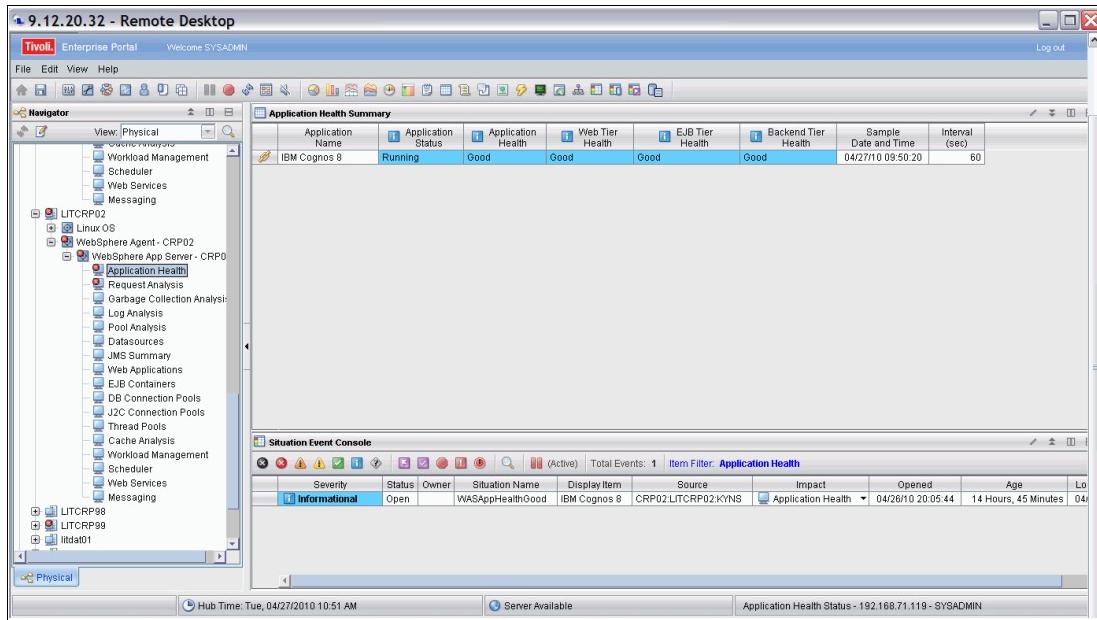
Figure 18-7   Tivoli Monitoring: WebSphere application or Cognos workspace

18.2 Monitoring cloud situations and events

A situation is a logical expression involving one or more system conditions. Situations monitor the condition of systems in the cloud.

All agents deployed in the cloud provide a set of predefined situations that monitor the performance and availability status of the monitored component and resources. These situations check for specific conditions and can trigger Critical, Warning, or Informational situation event indicators (also called alerts) in the Navigator. Some of the situations are set to start automatically. Other situations must be started manually.

When a situation triggers an alert, you can investigate the situation event by opening its workspace. If both a warning and a critical condition occur for the same workspace, the indicator always shows the highest-level situation event (the critical one).

Using predefined situations can improve the speed with which you can begin monitoring the cloud. You can examine and, if necessary, change the conditions or values that are being monitored by a predefined situation to those best suited to the cloud.

Figure 18-8 shows situations that are triggered in our Smart Analytics Cloud environment.
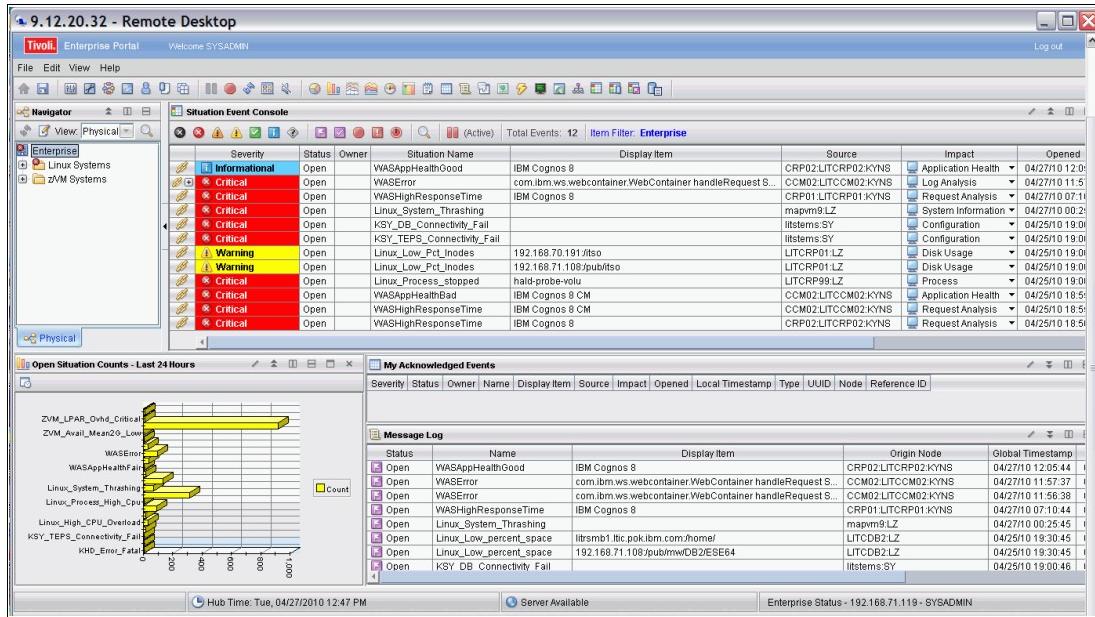
Figure 18-8   Tivoli Monitoring: Situations

19

Metering and billing

Financial Management considerations are fundamental for a profitable cloud service delivery and cost transparency. You get charged for what you consume. When implementing metering and billing, several common questions must be addressed:

► How do we allocate costs of shared IT resources?
► What do we charge for these cloud services?
► What usage do we measure?
► What is our cost do deliver cloud services?
► How do we bill users for using the cloud services?
► How do we automate the entire billing?

Usage-based pricing in cloud delivery requires knowledge of the service usage, resource usage, and delivery cost. Price model, billing plan, and metering plan are defined and set up at the early stage of the cloud service life-cycle during the service definition stage and service offering stage.

The metering starts when the resources are provisioned and continues until the service is terminated. Resource usage data is collected to determine the costs and the service usage data is collected for billing. Usage reports and costs reports are created regularly during the production period for charge back or capacity planning purposes. At the end of each period, the charges for using services and invoices are generated.

IBM Tivoli Usage and Accounting Manager (TUAM) helps to reduce IT cost by determining the cost of providing IT services. TUAM accurately assesses the usage of shared computing resources. Using it you can understand, track, and allocate the costs and invoice by department, user, and other criteria. TUAM transforms raw IT data into business information for cost allocation that covers business units, cost centers, applications, and users. TUAM delivers detailed information and reports about the intricate use of shared resources as it simultaneously masks the underlying complexity. Armed with this information, organizations can make informed decisions about where to reduce IT cost.

19.1 Metering and billing using IBM Tivoli Usage and Accounting Manager

Using IBM Tivoli Usage and Accounting Manager, we now know who consumes which IT resources, and we understand the resource consumption across multiple dimensions.

Businesses can determine the cost of those resources, including those that are shared in the cloud, and assigns cost to resources usage.

Cloud delivery organization can apply and allocate cost for chargeback, ROI, costing analysis, and billing that associates usage costs to consumers of IT resources.

Figure 19-1 on page 281 shows a sample user database usage report.

Figure 19-1   Sample Database usage report

Figure 19-2 on page 282 shows a sample IT expense report.
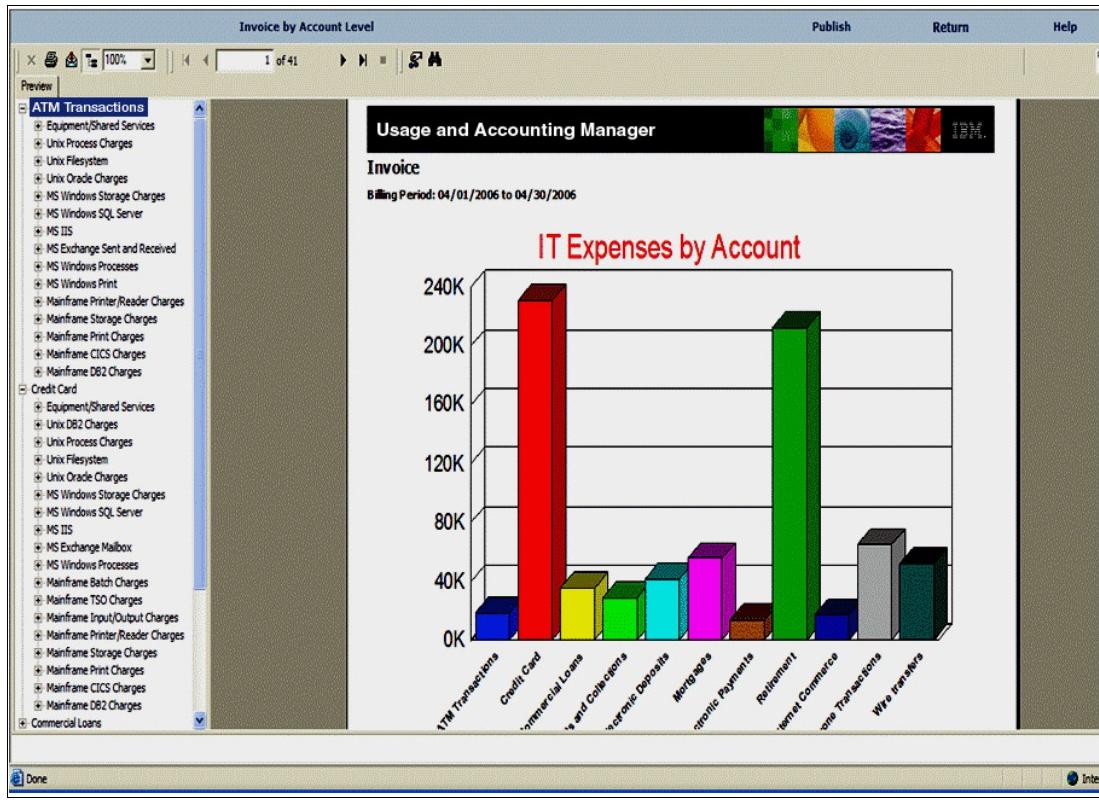
Figure 19-2   Sample IT expense report

Figure 19-3 on page 283 shows a sample cloud service billing report.

Figure 19-3   Sample cloud service billing report

20

Scenario: The cloud in action

In this chapter, we describe how all components and installations that we documented in previous chapters of this book are put together so that the end user can experience a Smart Analytics Cloud.

A corporation has a large multi disciplinary sales team and a finance team that both look for an analytics solution. They need it in a short amount of time; therefore, ordering and installing hardware and software and training their IT staff to install and administer it is not the answer. Their company just implemented a private Smart Analytics Cloud.

Both teams decide to use the functionality provided in the Smart Analytics Cloud.

The service that is provided with the Smart Analytics Cloud is defined in the service catalog (the service description is similar to the one provided in Chapter 4, "Scope of the Smart Analytics Cloud" on page 27). An onboarding application is also available with the functionality, as described in 6.2, "Onboarding application" on page 66.

Both teams will follow the same process:

1. A person from the team logs on to the onboarding application. They select the service they want to use from the displayed service catalog and type in the information that is required for onboarding, such as:

- Date when the Cognos Components are needed
- Environments that are needed: production, development, test
- Number of registered users
- User roles (Report Administrator, Report User, etc.)
- Data sources that need to be accessed with Cognos
- Education needed for using Cognos

2. The manager of the person sending the requests approves the application.

3. The Cloud service manager receives the request and checks whether further information is required and approves the request. This information is sent to the person who sent the application.

4. The Cloud Manager contacts the Cloud Administrator and the Cloud Consultant. Together they plan the required steps to add the departments to the cloud service (The onboarding process is described in detail in Chapter 16, "Onboarding" on page 253).

   - They check whether enough licences are available for the number of users that shall be added to the cloud. If the number of licenses is not available, the Cloud Manager contacts the procurement department to purchase additional licenses.

   - The Cloud Administrator checks whether enough storage space and memory is available. If further storage or memory is required, the procurement department is contacted.

   - The Cloud Administrator checks when free administrator resources are available to proceed with the provisioning and onboarding and sets the schedule for the provisioning activities.

   - The Cloud Consultants defines the tasks that must be done from the consulting side, such as giving support for creating new reports to the requesting department, or education in using Cognos, and so on.

5. After the planning is complete, the Cloud Administrator defines which parts of software must be cloned. The required hardware is set up (in our examples new Linux guests are created). Then provisioning starts. (The details are described in Chapter 13, "Provisioning and resource reallocation" on page 213 and Chapter 17, "Provisioning and resource management" on page 261. It includes the Cognos components with the underlying DB2 databases and clients and also the monitoring and systems management components.

6. After the provisioning is finished and tested the new users are added to the Cognos environment. The service requestor is informed that the new environment is set up.

7. After the environment has been set up and has been handed over to the users the regular operation tasks start. The system is monitored whether everything works as expected, patches to the system are applied in the

defined maintenance windows (this is shown in Chapter 14, "Tivoli Monitoring agent implementation" on page 221, Chapter 8, "Cloud Management architecture" on page 97, and Chapter 18, "Monitoring" on page 269).

8. At the end of each month the cloud manager creates an invoice that is sent to the responsible manager on the requestor side. The invoice is based on the consumed amount. Chapter 19, "Metering and billing" on page 279 describes the functionality in further detail.

These are the steps that the teams must follow when they want to use the services that a cloud provides.

How does this cloud solution differ from a regular solution?

In a traditional solution:

► Each team must set up a separate project with the IT department.

► An automated interface as the onboarding application does not exist, and the required information to plan the project is collected every time anew.

► An evaluation design about which hardware and software to use is made.

► The designs that each team makes are probably different.

► Hardware does not have to be purchased individually and set up manually.

► It probably needs time until it is delivered.

► The software is manually installed.

► Automation is not used in the same extent because each installation is rather a one of a kind implementation.

► Setting up development and test environments uses the same manual steps.

► Errors can occur during this process because the quality depends on the people conducting these steps.

► The monitoring, after it is manually set up, works the same way that the monitoring in the cloud works.

► The billing is probably not based on the resources used but is based on the price of the purchased hardware and software.

► In addition the IT department or the business department must educate two people for each solution to do the administration of the solution.

In summary, the Smart Analytics Cloud solution provides an accelerated solution at a reduced cost. The teams entering the solution face a standardized procedure for applying to enter the environment. As automation proceeds, you can even imagine that the entire provisioning is triggered and fulfilled automatically. The teams can use a preconfigured solution and configuration (including application and systems management) that saves them time in design.

Implementation time is reduced using automated provisioning. At the same time the quality of the provided solution was increased because tested components were implemented in a tested and proved way. The company of the teams saved money when purchasing the licences because it negotiated a larger number of licences with a single provider instead of talking to several suppliers over a small number of licences.

By using all of these standards and predefined services the teams however, remained flexible and independent in the types of reports they wanted to create and use.

A

WebSphere Application Server post-provisioning scripts

This appendix contains the WebSphere Application Server post-provisioning scripts used after the Cognos Report server was cloned.

289

A.1 Driving script

This is the driving script that is used on the cloned image. It must be run on the cloned image for LITCRP01, LITCIHS1, and LITCCM01. Example A-1 shows the sample was_post_clone.sh that we used in our test environment. It will modify the target wsadmin script for the target image before executing it.

> **Note:** The sample used here assumes that the source clone is LITCRPxx, LITCCMxx, or LITCIHxx.

Example A-1   Sample was_post_clone.sh

```
#!/bin/sh
#
# THIS SCRIPT WILL EDIT THE TARGET wsadmin SCRIPT WITH THE TARGET
HOSTNAME
# AND EXECUTE IT
#
if [ $# = 0 ] || [ $1 != "YES" ]
 then
      go="PROMPT"
 else
      go=$1
fi

echo $go

curHost=`hostname`

hostname_pfx=`echo ${curHost:0:6}`
#echo $hostname_pfx

case $hostname_pfx in
     "LITCRP") echo "You are working on the Report Server clone"
               sed -e "s!hostname_of_the_clone!$curHost!g"
/itso/scripts/clone_RP.sh > /itso/scripts/exe_clone_RP.sh
               chmod 755 /itso/scripts/*.sh
               . /itso/scripts/exe_clone_RP.sh $go;;
     "LITCCM") echo "You are working on the Content Mgr Server clone"
               sed -e "s!hostname_of_the_clone!$curHost!g"
/itso/scripts/clone_CM.sh > /itso/scripts/exe_clone_CM.sh
               chmod 755 /itso/scripts/*.sh
               . /itso/scripts/exe_clone_CM.sh $go;;
```

```
     "LITCIH") echo "You are working on the Gateway Server cloneServer
clone"
             sed -e "s!hostname_of_the_clone!$curHost!g"
/itso/scripts/clone_IHS.sh > /itso/scripts/exe_clone_IHS.sh
             sed -e "s!hostname_of_the_clone!$curHost!g"
/itso/scripts/create_web_server.sh >
/itso/scripts/exe_create_web_server.sh
             chmod 755 /itso/scripts/*.sh
             . /itso/scripts/exe_clone_IHS.sh $go
             . /itso/scripts/exe_create_web_server.sh $go;;
     *) echo "This script will do RP, IHS and CM server only"

esac

exit
```

A.2 Report server script

This script is executed by was_post_clone.sh, as a shown in Example A-1 on
page 290, for the Cognos Report server. The host name variable is substituted
by the driving script before execution. It must be run on the cloned images for
LITCRP01. Example A-2 shows the sample for the Cognos Report server.

The same script can be used for the Cognos Gateway server and the Cognos
Content Manager server by changing the corresponding server name and cluster
name inside the script. In our test environment, we created clone_IHS.sh and
clone_CM.sh from using this script respectively.

Example A-2   Sample clone_RP.sh

```
#!/bin/sh
# THIS SCRIPT WILL BE USED AFTER THE COGNOS REPORT SERVER WAS CLONED
# START BY SCRIPT was_post_clone.sh
#
# NOTE:
#  variable hostname_of_the_clone will be updated by was_post_clone.sh
#
if [ $1 = "YES" ]
 then
     your_ans=$1
 else
     echo "Are you sure to run this script (YES/NO)?"
     read your_ans
```

```
fi

aa=`echo $your_ans | tr '[:upper:]' '[:lower:]'`
currentPATH=`pwd`
instPATH=/opt/IBM/WebSphere/AppServer/bin
exePATH=/opt/IBM/WebSphere/AppServer/profiles/Appsrv01/bin

if [ $aa == "yes" ] || [ $aa == "y" ]
then
  cd $instPATH
  . manageprofiles.sh -create -profileName Appsrv01 -profilePath
"/opt/IBM/WebSphere/AppServer/profiles/Appsrv01" -templatePath
"/opt/IBM/WebSphere/AppServer/profileTemplates/managed"

  cd $exePATH
  . addNode.sh 192.168.71.97 8879 -username wasadmin2 -password
cogn0s00
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c 'AdminNodeManagement.listNodes()'

  cp /itso/scripts/soap.client.props
/opt/IBM/WebSphere/AppServer/profiles/Appsrv01/properties/
  cp /itso/scripts/ssl.client.props
/opt/IBM/WebSphere/AppServer/profiles/Appsrv01/properties/

  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c 'AdminNodeManagement.syncNode("hostname_of_the_cloneNode01")'
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c 'AdminTask.createSSLConfigGroup("[-name
hostname_of_the_cloneNode01 -scopeName
(cell):LITWSND1Cell01:(node):hostname_of_the_cloneNode01 -direction
inbound -certificateAlias '"'websphere server key'"'
-sslConfigAliasName ITSO_WAS_SSL_Settings -sslConfigScopeName
(cell):LITWSND1Cell01 ]")'
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c 'AdminTask.deleteSSLConfigGroup("[-name
hostname_of_the_cloneNode01 -scopeName
(cell):LITWSND1Cell01:(node):hostname_of_the_cloneNode01 -direction
inbound ]")'
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c 'AdminNodeManagement.syncNode("hostname_of_the_cloneNode01")'
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c
'AdminNodeManagement.restartNodeAgent("hostname_of_the_cloneNode01")'
```

```
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c
'AdminClusterManagement.createClusterMember("Cognos_report_cluster",
"hostname_of_the_cloneNode01", "Cognos_Report_server")'
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c 'AdminTask.setJVMMaxHeapSize("-serverName Cognos_Report_server
-nodeName hostname_of_the_cloneNode01 -maximumHeapSize 768")'
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c
'AdminConfig.modify("(cells/LITWSND1Cell01/nodes/hostname_of_the_cloneN
ode01/servers/Cognos_Report_server|server.xml#Property_1272495940478)",
"[ [name '"'LD_LIBRARY_PATH'"'] [value
'"'/opt/IBM/ITM/ls3263/yn/wasdc/7.1/toolkit/lib/ls3266:/opt/IBM/ITM/ls3
263/yn/wasdc/7.1/toolkit/lib/ls3266/ttapi:/home/cogdb2/sqllib/lib64:/ho
me/cogdb2/sqllib/lib32:/opt/cognos/c8_64/bin64::/opt/cognos/c8_64/bin:/
opt/cognos/c8_64/cgi-bin:'"'] ]")'
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c 'AdminNodeManagement.syncNode("hostname_of_the_cloneNode01")'
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c
'AdminServerManagement.startSingleServer("hostname_of_the_cloneNode01",
"Cognos_Report_server")'

echo Show server information
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c
'AdminServerManagement.showServerInfo("hostname_of_the_cloneNode01",
"Cognos_Report_server")'

  cd $currentPATH

else
  echo "Your answer is "$aa", script ended"
  exit 8
fi
```

A.3 Web server script

The web server script is executed by the was_post_clone.sh, as shown in Example A-3 after the host name variable is substituted, and it must be run on the cloned images for LITCIHS1.

Example A-3   Sample create_web_server.sh

```
#!/bin/sh
# THIS SCRIPT WILL BE USED AFTER THE COGNOS CONTENT GATEWAY SERVER WAS
CLONED
# START BY SCRIPT was_post_clone.sh
#
#NOTE:
#  variable hostname_of_the_clone will be updated by was_post_clone.sh
#
if [ $1 = "YES" ]
 then
      your_ans=$1
 else
      echo "Are you sure to run this script (YES/NO)?"
      read your_ans
fi

aa=`echo $your_ans | tr '[:upper:]' '[:lower:]'`
currentPATH=`pwd`
exePATH=/opt/IBM/WebSphere/AppServer/profiles/Appsrv01/bin

if [ $aa == "yes" ] || [ $aa == "y" ]
then
  cp /opt/IBM/HTTPServer/conf/httpd.conf /itso/scripts/
  cp /opt/IBM/HTTPServer/conf/admin.conf /itso/scripts/
  cd $exePATH

  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c 'AdminTask.createUnmanagedNode("[-nodeName
Unmanaged_node_hostname_of_the_clone -hostName hostname_of_the_clone
-nodeOperatingSystem linux]")'
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c
'AdminTask.createWebServer("Unmanaged_node_hostname_of_the_clone",
"[-name WebServer_hostname_of_the_clone -templateName IHS -serverConfig
[-webPort 80 -serviceName -webInstallRoot /opt/IBM/HTTPServer
-webProtocol HTTP -configurationFile -errorLogfile -accessLogfile
-pluginInstallRoot /opt/IBM/HTTPServer/Plugins -webAppMapping ALL]
```

```
    -remoteServerConfig[-adminPort 8008 -adminUserID admin -adminPasswd
cognos -adminProtocol HTTP]]")'
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c 'AdminConfig.save()'

  echo Show server information
  . wsadmin.sh -lang jython -conntype SOAP -host 192.168.71.97 -port
8879 -c
'AdminServerManagement.showServerInfo("hostname_of_the_cloneNode01",
"WebServer_hostname_of_the_clone")'

  sed -e 's!LITCIHS1!hostname_of_the_clone!g' /itso/scripts/httpd.conf
> /opt/IBM/HTTPServer/conf/httpd.conf
  sed -e 's!LITCIHS1!hostname_of_the_clone!g' /itso/scripts/admin.conf
> /opt/IBM/HTTPServer/conf/admin.conf
  cd $currentPATH

else
  echo "Your answer is "$aa", script ended"
  exit 8
fi
```

A.4 Cognos post-provisioning script

This script generates a configuration file and runs the Cognos configuration in
silent mode on the new clone. Example A-4 shows a sample Cognos post
provisioning script.

Example A-4   Sample Cognos post provisioning script

```
#!/bin/ksh
# Script runs cogocnfig.sh in silent mode to configure Cognos on a new
clone

if [ $# -lt 1 ]; then
        echo "usage: `basename $0` <new server ip address>"
        exit 1
fi

new_rpt_ip=$1
port="9080"
# new rpt server ip and port in format 192.168.71.52:9080
```

```
new_ip_port=$new_rpt_ip:$port

install_path=/opt/cognos/c8_64
script_dir=/cognos
rpt_cfg_tmplt=rpt_cfg.xml
new_cfg=rpt_cfg_new.xml
replace_string=rpt_ip_port
bkup_dir=/cognos

# Backup cogstartup.xml
echo "copying cogstartup.xml to $bkup_dir"
cp $install_path/configuration/cogstartup.xml $bkup_dir/.

# Backup Cognos cryptography files
echo "Saving encryption keys to $bkup_dir/configuration"
 $script_dir/backup_crypto_keys.ksh $install_path $bkup_dir

# Insert new report server IP and port into cfg file
echo "Creating new config file with new server ip"
sed "s/$replace_string/$new_ip_port/g" $script_dir/$rpt_cfg_tmplt >
$new_cfg

# Copy new config file into cogstartup.xml
echo "Copying new cfg file to
$install_path/configuration/cogstartup.xml"
cp $new_cfg $install_path/configuration/cogstartup.xml

# Run silent config
echo "Running cogconfig in silent mode"
echo " executing $install_path/bin64/cogconfig.sh -s"

$install_path/bin64/cogconfig.sh -s
```

Example A-5 shows the sample backup_crypto_keys.kshl.

Example A-5   Sample backup_crypto_keys.kshl

```
#!/bin/ksh

myname=`basename ${0}`

typeset -u p1=$1
```

```
if [ "$p1" = "HELP" ]
then
  echo
  echo "Usage $0 <cognos location> <backup location> "
  echo
  exit 0
fi

if [ $# != 2 ]
then
  echo "Usage $0 <cognos location> <backup location> "
  exit 2
fi

cognos_loc=$1
backup_loc=$2
ts=`date +%Y%m%d%H%M%S`
backup_dir=$backup_loc/configuration_$ts

if [[ ! -e $backup_dir ]]; then
  mkdir $backup_dir
fi
if [[ ! -e $backup_dir/csk ]]; then
  mkdir $backup_dir/csk
fi
if [[ ! -e $backup_dir/encryptkeypair ]]; then
  mkdir $backup_dir/encryptkeypair
fi
if [[ ! -e $backup_dir/signkeypair ]]; then
  mkdir $backup_dir/signkeypair
fi

if [[  -f $cognos_loc/configuration/cogstartup.xml ]]; then
  cp -p $cognos_loc/configuration/cogstartup.xml $backup_dir/.
fi
if [[  -f $cognos_loc/configuration/caSeriall ]]; then
  cp -p $cognos_loc/configuration/caSerial $backup_dir/.
fi
if [[  -f $cognos_loc/configuration/cogconfig.prefs  ]]; then
  cp -p $cognos_loc/configuration/cogconfig.prefs $backup_dir/.
fi
if [[  -f $cognos_loc/configuration/coglocale.xml  ]]; then
  cp -p $cognos_loc/configuration/coglocale.xml $backup_dir/.
fi
```

```
cp -p $cognos_loc/configuration/csk/* $backup_dir/csk/.
cp -p $cognos_loc/configuration/encryptkeypair/*
$backup_dir/encryptkeypair/.
cp -p $cognos_loc/configuration/signkeypair/* $backup_dir/signkeypair/.

exit
```

Competency centers: Sustained success through operational efficiency

While technology has always held the potential to transform business, its ability to accomplish that is often hampered by an organization's internal obstacles to adoption. The organizational complexity and the urgency of immediate need resulted in stove pipe implementations of Business Intelligence, Performance Management, and Data Warehouses solutions. This has resulted in distributed skills and overall inconsistency in the management, delivery and fulfillment of technology across the enterprise.

Organizations suffer from:

► Disconnected silos of data that reduce confidence in the information and drive inconsistency across the enterprise

► Lack of technology usage standards and best practices, which results in users seeking their own solutions

► Insufficient training and support resulting in wasted resources, missed deadlines, and end user dissatisfaction

► Misalignment between technology projects and business needs whereby incomplete or missing roadmaps and inconsistent executive sponsorship

hampers the long-term Information Management success and reinforces institutionalized, inefficient work arounds.

Only by removing common business obstacles can organizations realize the full value from investments in technologies like Business Intelligence and Analytics.

Figure B-1 shows the typical barriers to information management success.

Typical Barriers to Information Management Success

Culture
- Lack of senior management commitment
- Low trust in the data
- Communication gap between IT, Finance, LOB
- Lack of accountability & ownership for KPIs
- Data ownership conflicts
- Lack of demonstrated ROI

Process
- Projects not aligned with goals & strategy
- Disconnected business management processes
- Data quality is not improving
- Compliance issues, governance, risk
- Best practices not shared, accessible
- Missed deadlines, missed requirements
- No/poor/inconsistent methodologies & standards
- Lack of adequate help desk support

People
- Competing priorities, silos of activity
- Pockets of expertise that are not well-leveraged
- Business demand exceeds IT supply
- IT as a bottleneck
- Lack of resources – no business case
- Lack of training & skills
- No awareness of success in solving business pains

Data, Technology & Infrastructure
- Silos of data, poor integration
- Structured data vs unstructured data
- Response times do not meet expectations
- Many different tools, expensive to support & license
- Inconsistent meta data, data definitions, security
- Perceived to be difficult to use, hard to learn

Figure B-1   Typical barriers to Information Management success

More than ever, access to data and turning that data into insightful information is critical for competitive business in today's marketplace. The market dynamics, evolutions in technology and the continued rise in globalization, place enormous pressures and impose new challenges to organizations as they strive for efficiency and agility. This is further exacerbated by the ongoing growth in complexity and volume of data/information and compliance and regulatory requirements imposed on business today.

In response, companies continue to put Business Intelligence and Analytics on their list of top spending priorities to do more with less, improve higher quality services, and enable themselves to be agile in today's competitive marketplace.

IBM research indicates that most deployments tend to be tactical or localized at the department level, approximately 70% of the market approach today. Information has always been critical for business. In a fast changing global economy, timely access to accurate company data is increasingly essential for agility and growth. The reality is that today's organizations are often ill prepared from an operational and organizational perspective to reach the fullest rewards of their technology investments. Their individual collective potential remains unrealized if traditional technology implementation processes are followed. This is becoming a further challenge for companies today as they witness the information explosion occurring due to the increasing in the volume of digital data, the variety of information (data growth in unstructured content, emails, images, audio, video), and the velocity of decision making that is needed to support the rapidly changing business climate.

The typical organization has a complex and often siloed infrastructure in place that has burdened IT with overwhelming management needs, which are further exaggerated by increasing user requirements, and external influences for compliance and regulation. These systems and the appropriate processes to maintain and manage them have had few standards in place.

This complex environment, illustrated in Figure B-2 on page 302, boasts multiple tools, processes, applications, platforms and data silos with many users, who have different demands and roles, requiring access to the most current data that they can use to form decisions. This state of uncontrolled behavior leads to higher costs of deployment and management of the technology, a lack of user satisfaction, mistrust in the data, and often, mistrust in the capabilities of peers in IT or business to fulfill those needs. To gain the highest value from information as a strategic asset, organizations are striving to become operationally efficient and business effective, yet they are burdened by the history of their business model and inhibiting culture.
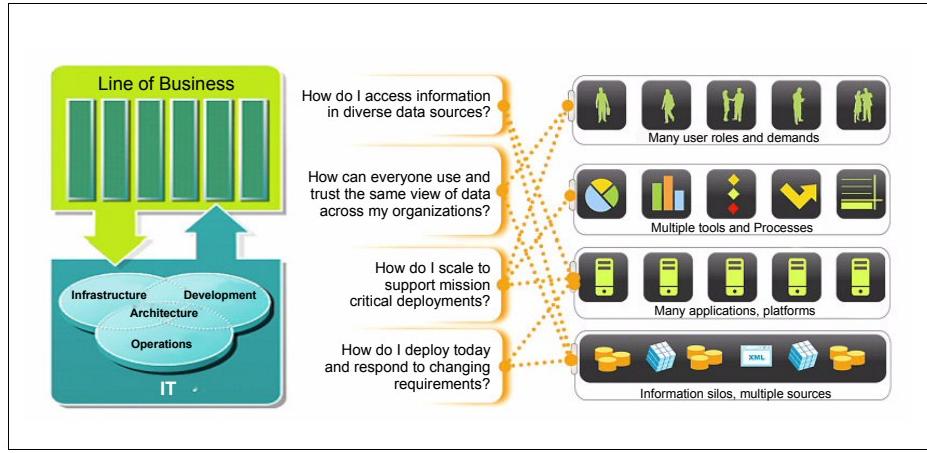
Figure B-2   The complex IT environment

With the drive toward collaboration on a larger scale, organizations need something more flexible: an Enterprise Information System that everyone can use. It consists of:

▶ An information-driven strategy to enable business priorities that align to corporate strategy

▶ An enablement roadmap and discipline that aligns people, process, and technology with this strategy

▶ A deployment plan that leverages technology best practices

▶ A foundation for trusted information that is governed for consistency and accuracy

With the advent of open Web Services architectures and innovations in Cloud computing technology, such scalable systems are now available. The challenge is in ensuring that it is adopted by the broadest group of users across the business, and by means in which the organizations is able to properly manage, maintain and support the scaling out of the technology investment. The challenge, all of this within a limited IT budget and with shortage of skills (see Figure B-3 on page 303 for business challenges). Seems like a difficult task but studies show that organizations that take a well managed, pragmatic approach to technology are able to unlock the value of business information. They do so, by leveraging Cloud based environments and by supporting that investment by process and people disciplines focussed on economy of scale and reuse of the technical and human assets in-hand. Organizations are creating working teams of IT and Business users, often commonly known as Competency Centers, to jointly work to improve the strategic value of information and to manage the

roll-out of enterprise wide Business Intelligence, Financial Performance Management and Predicative Analysis.

A Competency Center helps organizations define and establish the knowledge, standards and resources to support the organizations long-term information agenda for long term sustained success.

There is a greater need for IT to help address business challenges

Doing more with less
Reduce capital expenditures and operational expenses

Reducing risk
Ensure the right levels of security and resiliency across all business data and processes

Higher quality services
Improve quality of services and deliver new services that help the business grow and reduce costs

Breakthrough agility
Increase ability to quickly deliver new services to capitalize on opportunities while containing costs and managing risk

Figure B-3   Business challenges

It plays a critical role in how data is gathered, organized, accessed, and disseminated because it:

► Maximizes the efficiency, use and quality of Business Intelligence & Analytics across all lines of business

► Drives end user adoption by streamlining integration with routine workflows

► Encourages alignment between IT and Lines of Business

► Reduces organizational overhead associated with maintaining data consistency and use

► Increases the success of deployments of technology by helping organizations deliver more value, at less cost and in less time

► Improves business agility and technology management, which in turn drives business efficiency and effectiveness.

B.1 The Partnership of Business and IT

To provide trusted information upon which decision-making can be based, organizations need to gain alignment between the Lines of Business and IT, and across departments and functions in the management, access and business use of their information.

The reality however suggests that most organizations are not yet getting reliable information fast enough. Our studies indicate that results are best achieved by managing a central infrastructure, facilitated by Web Services architectures and innovations, such as Cloud Computing, and by wrapping these technical infrastructures with a supporting people & process operational design model (Figure B-4 on page 305). This model must include training & education, communication & evangelism services to advise & consult on best applied use of technologies, and focus its attention on freeing up time spent supporting redundant applications so that the organization can become more agile and capable of responding to the business.

Recognizing the need for operational efficiency and business effectiveness, organizations are embracing strategic transformations through a Competency Center to:

► Mature Business Intelligence into Performance Management

► Develop a single unified Business model

► Reduce and Manage costs through Standardization

► Take advantage of best practices in the use and reuse of both technical and human assets

► Instill disciplines and practices for Data Governance

► Implement and promote measures of success and ROI

Doing so, allows organizations to focus on value creation in the Information Management process. Their business model under constant improvement to focus on enabling timely and accurate measurement, monitoring and management of key performance metrics by employees at all level of the enterprise, and by facilitating better decision making throughout the enterprise at the strategic, tactical and operational levels.
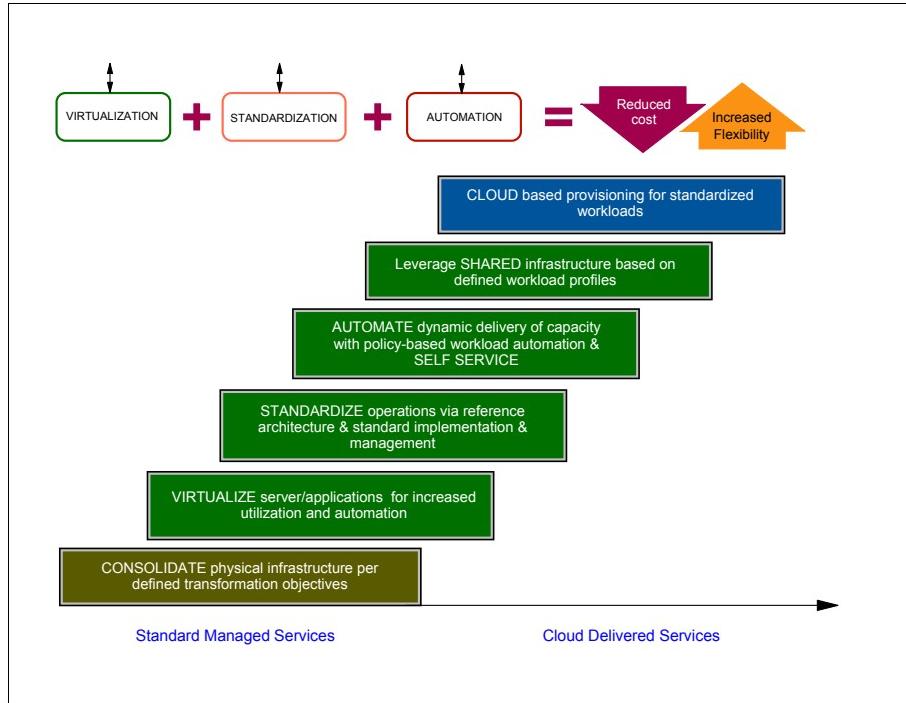
Figure B-4   IT capabilities provided as a service over an intranet, within the enterprise and behind the firewall

The Lines of Business and IT leaders need to work together to transform their business through Business Intelligence and Analytics to take advantage of these capabilities to the fullest. CIOS have an important role to play here. IT can facilitate the integration of geographic or departmental operations, manage individually owned processes through disciplines of governance, align and eliminate disparate silo sources into a centralized book of business. This helps ensure that the reliable data is available to users to meet the speed of business. Business Intelligence and efforts centered on harmonization of the infrastructure and standardization of the technologies and business processes to support them are key to the success of the Competency Center.

Together, working with the line of business and business leaders, IT can also develop common data definitions and policies, establish oversight and governance to ensure the success and scaling of standardization based initiatives. A formal interlock between Business and IT is absolutely critical. Business needs to access the trusted data and gain a common version of the truth that will aid in making business decisions. Their role is focussed on value creation and their knowledge of the function, line of business or task to turn the data into information for insight. IT needs to support the demands, while

maximizing and improving the technology foundation, and management thereof, by more efficiently using resources, skills and tools throughout the organization. This interlock of business and IT, in turn, will assist the organization in better managing costs, in realizing greater operational efficiencies to save money, and in operating more effectively to improve overall enterprise strategy.

B.1.1 Business Intelligence to Performance Management

Beyond Business Intelligence, Performance Management offers the next level of capability and management insight. It leverages Business Intelligence, and includes the critical linkages to the enterprise planning and control cycles, which support key management processes. Performance Management is about empowering decision makers throughout every level of the enterprise. A complete system not only offers analytics and reporting for critical business insight, it also provides the integrated planning capability to bridge evaluation with vision, strategy and execution.

B.1.2 A Unified Business Model for Information Management

Issues of fragmented data, disparate perspectives and manual systems can be overcome with a robust performance management system. Technology alone isn't enough however. A competency Center with the right charter and cultural design to compliment an organization is needed to ensure widespread adoption of technology through a unified managed approach. Since information has become a strategic asset, the Competency Center provides the foundation for the development of the strategic plans, priorities and means of leveraging human and technical assets to ensure that this valuable asset is leveraged to its entirety. If you have a management mandate without cultural acceptance, the result is half-hearted adoption by business users and poor outcomes. If you have the culture onside without a management mandate, the result is fragmented usage. Users pursue performance management 'in silo" without strong management direction and support, making it difficult to reconcile different views of the business. And without process, it becomes difficult to encourage and showcase a repeatable, consistent manner of managing, deployment and using the technology that the organization has invested in.

B.1.3 Formal mandate for standardization

Through standardization, organizations want to encourage the proactive use of information management technologies and to do so in a repeatable process that leverages the companies assets and skills, rather than re-inventing the wheel each time for each project. Consistent management, deployment and use of Business Intelligence enables an organization to be flexible and agile; to

empower users to make informed decisions that drive performance. Consistency is the result of having standards in place throughout the enterprise organization: standards for technology tooling and deployment, standards around how people work and interoperate, and standards that enable repeatable methods and guidelines of use. Standardization brings benefits to the whole organization. It simplifies the IT environment and can significantly reduce administrative efforts.

Standardization reduces system complexity by addressing the problems associated with managing and maintaining multiple tools, multiple versions of the environment. It simplifies and paves the way for Business Intelligences to be less of a tactical tool and more of a strategic solution.

IBM and analyst studies around the world, illustrate that most organizations continue to have a plethora of redundant technology tools. Companies continue to deploy new technologies as they evolve over time, and decentralized purchasing decisions are being made based on project-specific need versus company need. In addition, the aggressive market consolidations have further led to mergers & acquisitions, causing organizations to double the number of tools and technologies being supported almost overnight. Our IBM studies on Standardization clearly show that with standardization proclaimed and executed in an organization, companies are better prepared and receiving greater benefits from improved business processes, are seeing a reduction in enterprise operating costs, are attracting and retaining better customer relationships and are improving enterprise effectiveness.

The benefits from standardization are centered on enterprise efficiency and effectiveness gained through managing costs and increasing ROI:

► Increased transparency of information

► Reduction in overall TCO

► Consistent / repeatable delivery of technology

► Business User Satisfaction

► Utilization of skilled individuals / decreased staffing costs through reuse of best practices

► IT efficiency and its ability to respond to the business quicker

► Increased end user adoption and self-service

► Increased decision making speed and accuracy

► Recognition of new ways to take advantage of information

Conversely, we can also look at the cost of not standardizing, where most companies agree that it then becomes almost impossible to have controls in place to manage the numerous Business Intelligence initiatives underway in an organization. It becomes too complex to drive forward under one business

strategy and the technologies remain tactical in nature rather than strategic. Without standardization, organizations share common, painful realities of:

► No consistent view of the data: Complexity and duplication of the environments and tooling, leading to wasted time and money spent on reconciling figures and disputing findings of business views.

► Diluted expertise and lack of skills: Companies need to retain higher number of staff to ensure that they have skills required for all of the different tools and overreliance on staff that have specific knowledge of business. Relying on different tools means that there is a diluted expertise in the technology, and each time a new tool is implemented, there is a feeling of reinventing the wheel.

► Tactical Approach: The business and IT want Business Intelligence to drive and support business strategy. However, without standardization, the Business Intelligence tools continue to be used at a tactical level for each department, rather than creating or following a higher level strategy. Fragmented Business Intelligence systems prevent organizations from cross-functional analysis, spotting trends, or making the most of information. This demands a common database, common process, and common Business Intelligence technology.

The right decision made at a tactical project level is often the wrong decision made when viewed from a more holistic, company-wide perspective. By focusing on a set of standards, the entire life cycle of information decision making can be improved since the burden of time and investment shift from supporting multiple tools, methods and people, to building upon a single operational framework that refines and improves a limited number of people, process and technology choices.

B.1.3.1 Principles of Data Management and Governance

Better decisions drive the bottom line, and are in turn supported by quality data. In this case, "quality" means not only accurate data, but consistently available, understandable and relevant data as well. When organizations lack confidence in their data, their decision-making capabilities suffer, which in turn affects organizational performance. Data Quality often suffers because the silo-based environment remains institutionally disconnected. Business units create their own reports against their own data sources, resulting in inconsistent information across the enterprise, and IT often does not always know where and how the data is being consumed to provide the highest level of support for transparency.

Data Quality Management includes the methods to measure, improve and certify the quality and integration of production, test and archival data. To remediate any significant data quality issues, data ownership must first be clearly defined and communicated across the organization and data owners held accountable for the

quality of their data. Because business is responsible for establishing the business rules that govern the data and are ultimately responsible for verifying the data quality, a strong partnership between business and technology is essential for any data quality management effort to succeed. Data quality issues must be identified, assessed and resolved as quickly as possible. Failure to adequately identify and resolve data quality issues results in a loss of confidence in (and funding of) the Information Management program.

Most organizations still lack a governance program today. Data Governance is a key function of a typical Business Intelligence Competency Center as organizations look to develop a managed, consistent approach to maintaining cohesive policies and process models.

B.1.3.2 Best Practice for use and reuse of both technical and human assets

Information based initiatives like Business Intelligence and Analytics rank high among strategic investments, but sometimes fail to deliver as promised. Implementations that take a technology-centric approach fall short because they are unable to address a given business need. Narrowly focussed projects often result in individual areas of an organization iterating their own solutions for making use of the new toolset. If best practices are develop at all, they are typically not shared across a broad enough cross-section of the organization to truly have impact. With each business area pursuing its own solutions, economies of scale are not realized, and support resources are unduly strained. To overcome these shortcoming and pitfalls, it is important that organizations create a mandate for standardization and enlist the support and enthusiasm of end users by:

► Knowing the organization: Find out how many technology tools are being used. Evaluate the costs of key workflows, such as financial reporting and budgeting, to find opportunities for improvement.

► Publishing early successes: When standardization has early wins, people become far more dedicated as champions, investors, and participants, if they are confident of the benefits to come.

► Gaining the support of senior management: Strengthen initiatives as a company wide endeavour, helping people overcome their resistance to change. A top-down approach is more successful for hierarchical cultures where management can set the timetable and measures of success. IBM studies indicate that executive sponsorship, whether in Business or IT, is the number one factor that can either make or break a Competency Center endeavour.

B.1.4 Measure return on investment

Return on investment (ROI) is important for gaining buy-in from management. Important that the measures of success align to your cultural metrics whether they be in measuring breadth (more participants equals higher adoption), repeatability, cost savings, improved collaboration, or measure of collective knowledge (reuse of information).

B.2 What is a Competency Center

By helping promote an Information-centric culture throughout the organization and by facilitating adoption in areas of the organization that have traditionally been bypassed, Competency Centers can help organizations unlock capabilities in their data and their people, and become more agile and competitive in the process. A Competency Center can provide the centralized knowledge and best practices to help make this broader technology initiative possible.

With this in mind, many organizations are creating working teams of IT and Business users to improve the strategic value and deployment of the technologies. Organizations today need to bridge the gap between Business and IT. In doing so, many have relied on methodologies and principles of a Company Center. Though some organizations are sensitive to the naming of competency centers as Business Intelligence Competency Centers (BICC), the basic fact remains that it is simply good business to pay attention to the operational business model of any company. Many organizations have established a Competency Center by nature of their business model design, but in most cases, the historical behaviors of companies have gone unchallenged.

In some organizations, the Competency center is a formal organized department, grouping people with interrelated disciplines, domains of knowledge, experiences and skills for the single purpose of promoting expertise throughout the company. If technology is to extend beyond tactical deployments to become a broader-based solution, a well managed, predictable approach to operating the day to day and recurring business cycles is needed. This has become increasingly more important today as the market dynamics and speed of decision making, competitive pressure & opportunity squeezing organizations to be highly transparent, proficient and effective in their daily operating & decision making.

Today, more than ever, there is a greater need for IT to help address business challenges to:

► Do more with less: Reduce capital expenditures and operational expenses

- ► Reduce Risk: Ensure the right levels of security and resiliency across all business data and processes

- ► Provide higher quality services: Improve quality of services and deliver new services that help the business grow and reduce cost

- ► Open new opportunities with breakthrough agility: Increase ability to quickly deliver new services to capitalize on opportunities while containing costs and managing the risk

The challenge however remains, most organizations maintain an inflexible infrastructure, complex operation and dynamic relationship between business and IT. The developments in Cloud technology are helping to address many of these challenges from a technical perspective, but optimized decision making extends far beyond the traditional IT boundaries, it extends in the to unique, historical and ever changing dynamics of people and process of the organization itself.

Let us not get hung up on the word itself. The BICC simply is the most common naming convention that evolved over the last decade. Others names exist, for example: Centers of Excellence, Central Service Report Factories, Center of Knowledge, and as the name implies, there is no one size design that will fit all. But if we look at the bare essence of the Competency Center methodology, we uncover the four key tenants that they all focus on:

- ► Harmonization of People and Process and technology: Deploying a consistent set of standards and process

- ► Economy of Scale: Leveraging best practices and domain knowledge for sharing, transferring and broadening skills

- ► Managed Coordination: Coupling disciplines of Governance to ensure not only data quality and consistency, but also ensuring that binding decisions and consensus recommendations help drive the company in one direction

- ► Partnership: Business users from multiple functions, working in partnership with IT to form a virtual community of knowledge

Organizations today are developing these enterprise-wide initiatives with standards around technology. The Cloud innovations today help accelerate this movement and scale to meet the growing demands of disperse options. The Competency Center supports this effort by providing the reusable knowledge, disciplines and best practices to help make the broader technology initiatives possible:

- ► Business Intelligence Competency Centers (BICC) has become more common around the domain of reporting, analysis, and dashboards.

- ► Integration Competency Centers are emerging around the discipline of extracting, cleansing, and transforming data to create enterprise-wide

coordinated data marts and data warehouses. This, together with Master Data Management has become increasingly vital to establish a foundation of trusted information.

► Data Governance programs are gaining more visibility as evidenced by the formation and success of the IBM Data Governance Council of leading firms whose interest is to advance the adoption of data governance around a clear framework and operational model.

► In the realm of Enterprise Content Management (ECM), more and more organizations are adopting ECM Centers of Excellence to better take advantage of investments in unstructured data, records management, and associated business process improvements.

► In the office of Finance, teams are being formed with a focus on Performance Management Centers for the executive and C-level users to drive an enterprise scorecard-driven perspective of historical performance and future plans, budgets, and targets with common key performance indicators (KPIs).

What is now emerging is the integration and alignment of all these various groups into a more coordinated, often virtual, team called an Information Management Competency Center, as illustrated in Figure B-5 on page 313. More than a just a set of centers, it is a well-defined and coordinated community of both business and IT stakeholders with authority on the direction and use of information assets.
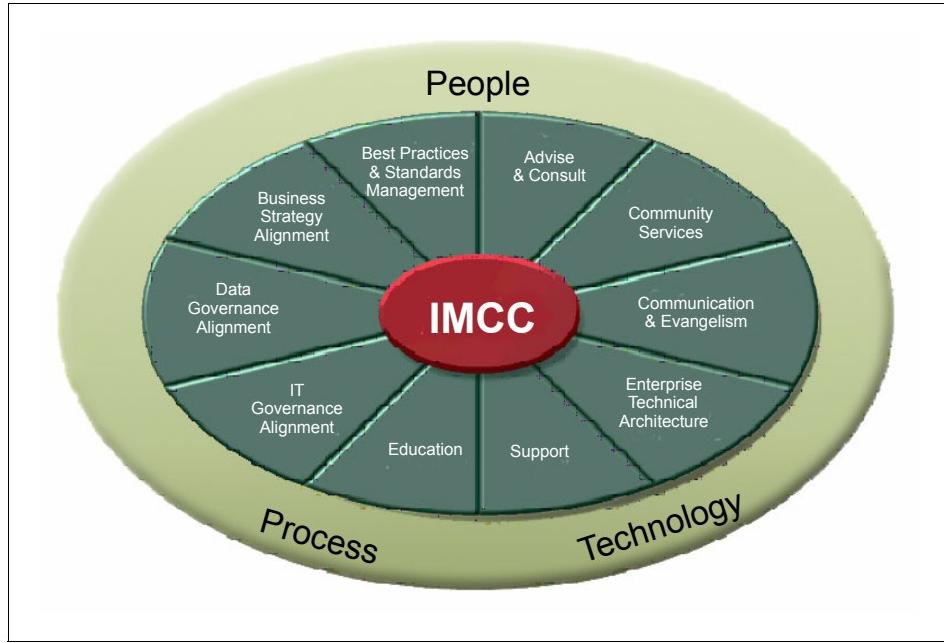
Figure B-5   The Information Management Competency Center model

As organizations realize that Business Intelligence, Performance Management, and Data Management are more than just tools, and as they look to break through the proliferating silos of data, they are investing resources to help their organization manage information more effectively, make better decisions as a result, and use that capability to gain competitive advantage. To approach Information Management components as strategic assets, many organizations are combining IT resources and Business Intelligence, Performance Management, and Data Management users into coordinated teams.

These emerging organizational structures bring together people - with interrelated disciplines, domains of knowledge, and experiences – with well defined processes and standards. Their goal: To achieve greater IT and business efficiencies on a foundation of trusted information which delivers more effective decision making across the enterprise.

The Competency Center enables repeatable, successful technology deployments through the development of people, technology and process - in ways that makes sense to an entire organization or division, rather than just a single project. Processes and procedures are often the most overlooked aspect of Business Intelligence deployment. These are critical components to sustaining success due to the critical importance of governing how data is accessed, developed and disseminated. This has become an even more important topic

over the past few years due to compliance and regulation. Processes and procedures help eliminate duplication of effort, inconsistent results and delays in dissemination.

While initial deployments might have focused on targeted uses, the current trend is to deploy Information Management more broadly across the organization. As part of this trend, Information Management (encompassing Business Intelligence, Analytics, Performance Management, Data Integration, Data Warehousing and Enterprise Content Management) is rapidly migrating from specialists, analysts, and knowledge workers toward executives and everyday business users looking for direct and faster access to the information they need to make better decisions and get their jobs done.

To maximize the value of having information in the hands of this new, wider audience, organizations are challenged with implementing, managing and supporting these tools and capabilities across divisions, regions and functions with a diverse set of user needs and skills. Processes that were already in place for conventional applications and platforms are no longer sufficient because the use of information is so dynamic with constantly changing requirements. Creating a Competency Center to strengthen the partnership between the line-of-business stakeholders and IT, to develop and communicate a clear Information Management strategy aligned with business strategy, to standardize technologies and processes, and to take advantage of reusable knowledge, disciplines, and best practices, is critical to realize the full value from Information Management based investments and drive long-term success.

Regardless of name, a Competency Center is an organizational structure that groups people with interrelated disciplines, domains of knowledge, experiences and skills, for the purpose of promoting expertise throughout an organization.

B.3 Value creation

To understand the need for a Competency Center, it is important to first understand the business problems and related Information Management challenges that many customers face.

First, Business Intelligence and Analytics initiatives rank high among strategic investments but sometimes fail to deliver as promised. Implementations that take a technology-centric approach fall short because they are unable to address a given business need. Narrowly-focused Business Intelligence and Analytics projects often result in individual areas within the organization iterating their own solutions for making use of the new toolset. If best practices are developed at all, they're typically not shared across a broad enough cross-section of the organization to truly have an impact. With each business area pursuing its own

solutions, economies of scale are not realized, and support resources are unduly strained.

Additionally, data quality also suffers because the silo-based environment remains institutionally disconnected. It lacks a governance program which traditionally develops and manages a consistent approach to maintaining cohesive policies and process models. The Business Intelligence and Analytics fail to take root to the degree that they should because the net benefits of broader implementation are not realized – which motivates end users to simply continue working as they have. As regulatory and competitive requirements increase the need for greater volumes of work, shorter timelines and increased decision-making precision, organizations that fail to improve the level of their game risk falling behind, expose themselves to regulatory breach, or both.

A Competency Center addresses this gap by focusing on the root causes of misalignment. By consolidating representation from key areas across the organization, it can encourage the kind of ongoing discussion that maximizes alignment. Under the single managed approach, cross-functional teams identify opportunities for developing, deploying and monitoring best practices and ensuring precise linkages between fast-changing business requirements and the tools that will address them. By driving business requirements through the heart of all Business Intelligence & Analytics based initiatives, a Competency Center is an organization's best means of maximizing business value. It does so by capturing latent opportunities in existing data stores, putting business-friendly tools into more hands at all levels of the organization, and measuring the progress on their information management journey. There is a clear relationship between those companies that are outperforming in their industry based on the working practices they have put into place to enable their wining conditions based on operational & organizational efficiencies.

B.3.1 Having confidence in the data

Better decisions drive the bottom line, and are in turn supported by quality data. In this case, "quality" means not only accurate data, but consistently available, understandable and relevant data as well. When organizations lack confidence in their data, their decision-making capabilities suffer – which in turn affects organizational performance.

Against that backdrop, companies clearly want to get the most out of their Business Intelligence, Performance Management, Analytics or Data Management investments.

And in survey after survey, our customers tell us the kind of obstacles they face in trying to get the full value from these investments, as shown in Figure B-6:

► Disconnected silos of data that reduce confidence in the numbers and drive inconsistency across the enterprise.

► Lack of technology usage standards and best practices, which results in users seeking their own solutions.

► Insufficient training and support resulting in wasted resources, missed deadlines and high rates of rejection by end users.

► Misalignment between technology projects and business needs. Incomplete or missing roadmaps and inconsistent executive sponsorship hamper long-term BI and PM success and reinforce institutionalized, inefficient work arounds.
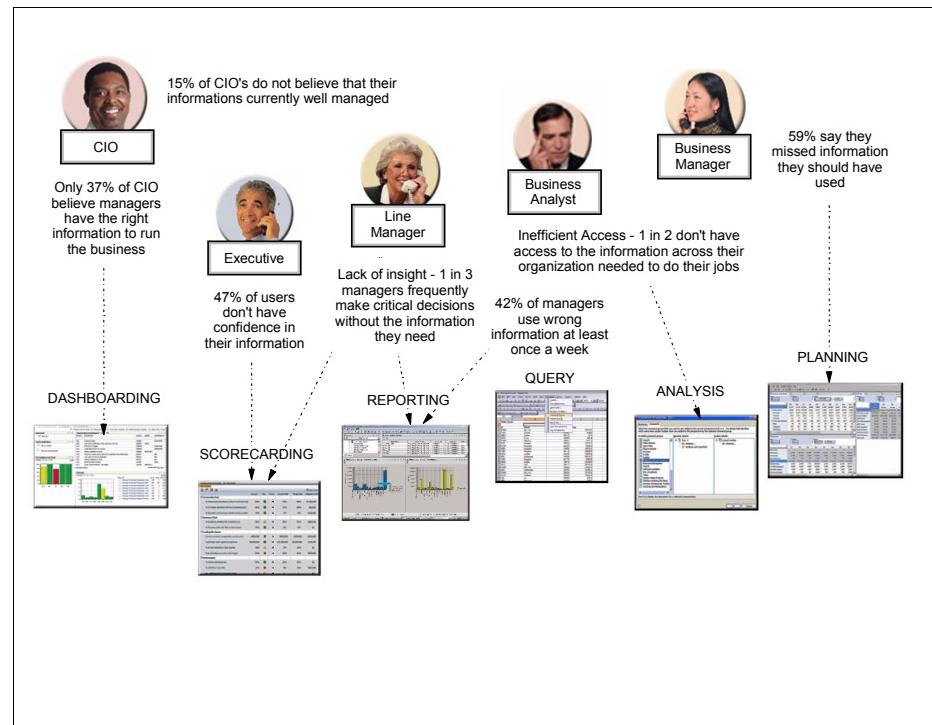
Figure B-6   Data confidence

To address these obstacles, organizations must have the following key elements in place:

► An information-driven strategy to enable business priorities that align to corporate strategy

► An enablement roadmap that aligns people, process and technology with this strategy

► A deployment plan that leverages Information Infrastructure best practices

► A foundation of trusted information that is governed for consistency and accuracy

The Competency Center provides the framework within which an organization can incrementally work through each of these steps to drive the transition from an application-based agenda to an information-based agenda, and in doing so unleash the latent potential inherent in the pools of data that automated systems have been generating for years.

As promising as Business Intelligence, Performance Management, Analytics, and Data Management can be to organizational performance, their individual and collective potential remains unrealized if traditional technology implementation processes are followed. Processes and standards are often the most overlooked deployment aspects. These are critical for sustained success due to the importance of governing how data is gathered, cleansed, organized, accessed and disseminated. This becomes even more vital with increased focus on compliance and risk.

If these technologies are to extend beyond tactical deployments to become broader-based solutions that support the organization's long-term information agenda, a managed, predictable, repeatable approach is needed. With this more strategic need in mind, a Competency Center can help organizations define the knowledge, standards and resources to make this happen. A Competency Center is essential to the strategic deployment of technology because it:

► Maximizes the efficiency, use and quality of BI, PM and DM across all lines of business.

► Drives end user adoption by streamlining integration with routine workflows.

► Encourages alignment between IT and business areas.

► Reduces organizational overhead associated with maintaining data consistency and use.

► Increases the success of deployments by helping them deliver more value, at less cost and in less time.

► Improves business agility and technology management, which in turn drives business efficiency.

In this age of competitive and regulatory pressures, businesses find themselves challenged as never before to exploit every opportunity to improve efficiency and effectiveness for competitive advantage. The Competency Center derives value from three aspects:

► IT efficiencies and cost savings derived from the economies of scale that are enabled

► Business efficiencies and cost savings derived from improved business processes

► Business effectiveness and value from better decision making derived as a result of having an established foundation

B.3.2 IT efficiencies

Typically, a Competency Center is the catalyst for centralizing infrastructure, such as servers and standardizing Business Intelligence, Performance Management, Analytics, and Data Management software, with a shared services approach. This might also include centralized data, though in some cases this might be left to various business divisions to maintain. Tenant LOB users will often "subscribe" to this service with potentially significant costs savings compared to maintaining separate silo implementations.

Shared service centers might also provide a central pool of talent for educating and supporting the LOB community of users. This helps maximize tool functionality and best practices adoption – which in turn drive end user adoption and higher levels of self-service by the business. By creating a common location for these Competency Center functions, IT leverages economies of scale with common education and support programs. This drives costs down and relieves IT of the traditional reporting backlog associated with processing such requests from across the organization.

As a design authority, the Competency Center typically takes responsibility for common Business Intelligence-related standards, such as naming, design, templates, reporting, and setting up and managing a Business Intelligence portal, among others. Beyond standards, the Competency Center also assumes accountability for methodologies, portfolio management, and projects deployment standards by participating in formation of data and information governance programs. These common standards make IT more efficient by re-using the standards rather than re-inventing them in different silos. Project planning and execution is streamlines because project managers can draw from pre-built standards that have already been proven.

The Competency Center is also often responsible for adherence to proper IT processes for security, version control, and production deployment, this reduces

compliance- and audit-related risks for IT and the organization at-large, and makes governance more consistent and efficient compared to a silo-based approach.

The Competency Center regularly communicates the results of successful projects, the status of the BI roadmap, and evangelizes how Business Intelligence and related Performance Management competencies will address specific business requirements. In this way, the Competency Center provides clarity on how IT is providing value to the business, not only from a cost savings perspective which results from operational efficiency, but also through value creation which results from business effectiveness.

Table B-1   IT benefits and the added value of Competency Centers

IT benefit	How Competency Centers add value
Optimized cost of ownership, breaking down silos	Drives shared, standardization Information Management architecture and software
Improved productivity, responsiveness, backlog	Drives user adoption, enables self-service
Improved compliance, audit, governance, mitigate risks	Ensures adherence to IT processes, policies, and standards
IT investments bringing business value and highest enterprise ROI/return	Communications and recognition of success
Innovating new ways to use information and maintain a competitive edge	Research information technology trends, advise on new opportunities to add value

B.3.3 Business efficiencies and effectiveness

Investments in Information management based technologies are only valuable if they are tied to business priorities. The Competency Center provides the critical link by which business strategy is translated into actionable information. It is the focal point for consensus and decisions on sometimes competing priorities and thus determines the Information Management roadmap as it evolves over time. This ensures proper alignment with strategy and fosters effective and efficient collaboration between the business and IT.

The business must trust the data in order to make effective decisions with confidence. This requires appropriate stewardship to continually improve the quality of the data. The Competency Center must be aligned with existing Data Governance programs by ensuring both the business processes and IT

processes – and in particular the data warehouse/integration processes – are in place and adhered to for data quality assessment and improvement. In some cases, the Competency Center might even go a step beyond alignment, and might actually be responsible for such programs because many organizations still today lack a formal Data Governance framework with stewardship assignment and accountability by role. Competency Centers are also often the source for reporting on quality and process-related metrics for proper and effective data governance. This in turn mitigates the risks associated with poor, incomplete, or inaccurate data, as it provides a performance baseline for measuring effectiveness, and an opportunity to ramp performance over time.

As mentioned previously, the Competency Center provides a central service for education, support, and best practices. Also, with knowledge of the data and experience in advanced analytics, the center can advise and consult for the business to drive more self-service out to the LOB knowledge workers and power users, albeit, in a controlled, governed manner. The value to the business is a greater degree of responsiveness and agility to rapidly changing business requirements, without having to go through an IT bottleneck. The business can now focus on improving specific business management processes by making them more efficient – for example, reducing reliance on time consuming Excel maintenance and verification, or enabling more effective processes that were previously starved for information.

With silos of information, it is common to have the same data defined in different ways, or for different data to be mistaken for the same definition. This is often most visible to the business at the KPI/metrics level where data is exposed on scorecards, dashboards and reports. To resolve issues of definition requires clear ownership and accountability, for example, who owns revenue/head count? How should that be defined? How does that drive alignment from strategy to execution down through the organization?

The Competency Center can be the catalyst and facilitator to help resolve questions of ownership, definition, and consistency in the data. This can then be reflected in the deployed Information Management environment through features, such as business glossaries and data lineage displays. This in turn ensures everyone in the organization can be more effective and efficient because everyone is working from the same playbook of data definitions and meaning. Increased and accelerated adoption of the complete Business Intelligence life cycle encourages the creation and use of a "trusted version of the truth" across the entire enterprise, increasing business confidence in the data being used to drive decisions.

By putting Business Intelligence and Analytics tools and processes into more hands throughout the organization, Competency Centers help drive greater levels of self service, which in turn improves user satisfaction levels and allows

more value-added deployment of organizational resources. IT support teams, for example, freed up from managing routine requests for reports from Business Intelligence-less business users, can reallocate resources to creating value elsewhere.

In this way, a Competency Center can help enforce more consistent Business Intelligence standards through registration and guidance, and identify new opportunities to take advantage of the toolset. This results in an alignment of technology to strategic goals and a clarity of vision for future coordinated Business Intelligence. It also creates both tangible and intangible value throughout the organization, including competitive differentiation and increased ability to routinely meet ever-evolving regulatory requirements.

Table B-2   Business benefits and the added value of Competency Centers

Business benefit	How Competency Centers add value
Better alignment with strategy, business, and IT collaboration	Drives Information management roadmap, building consensus on priorities and funding
Improved confidence in data, mitigate risks	Drives information quality & stewardship for accountability
Responsive to changing business environment, improved business management processes	Enables self-service, shared best practices, improved analytics across the enterprise
Clearer accountability for business results, align execution	Facilitates data ownership, common business glossary (KPI / metrics definition)

More strategically, a Competency Center also serves to educate key stakeholders about the advantages of employing Information Management technologies. It helps to build the lines of communication between regions and departments – and in doing so prevents a silo-driven approach to knowledge management. A Competency Center clearly demonstrates the value of these Information Management investments by illustrating throughout the organization the breadth, depth, completeness, accuracy and timeliness of information. This capability extends beyond mustering support for the next phase of investment to creating an information-centric culture that supports the organization's long-term strategic Information Management roadmap across the enterprise.

IBM and analyst studies show that Competency Center based organizations have better alignment between IT and the business, and experience more pervasive use of Business Intelligence across the organization. This helps them outperform organizations without a discipline from an alignment of Business and IT perspective, from an adoption perspective and from a business value creation perspective. The Competency center based organizations provide tremendous

value (tangible and intangible) to IT and business throughout the entire organization:

► Drives a lower total cost of ownership of Business Intelligence and technology solutions with reduced implementation costs or deployment risk through:

– The consolidation of best practice functions and services, allowing rapid, repeatable successes from other deployments.

– The centralization of competency and operational efficiency which maximizes the use of technology resources and assets.

– The ability to provide strategic BI deployment planning—accelerating rollout success.

– The management of healthy tension of project prioritization.

► Higher and faster adoption of the complete Information Management life cycle and trusted version of the truth across the entire enterprise which improves user satisfaction and self -service.

► The ability to enforce standards through registration and guidance and the ability to identify new opportunities to take advantage of technology. These results in an alignment of technology to strategic goals (competitive differentiation / regulatory requirements) and a clarity of vision for future coordinated Business Intelligence.

The Competency Center also serves to educate key stakeholders about the advantages of employing technology. It further helps to build the lines of communication between regions and departments to prevent a silo-driven approach to implementation. In doing so, the technology solution will clearly demonstrate its value through the breadth, depth, completeness, accuracy, timelines of information available throughout the enterprise.

Having trusted information requires appropriate stewardship to continually improve the quality of the data. The Competency Center should align with an existing Information Governance program (or sometimes should be responsible for such a program) by ensuring the right business and IT processes are in place and adhered to for data quality assessment and improvement. In addition, Competency Centers are often the source for reporting on quality and process metrics for proper and effective information governance which mitigates the risks associated with poor, incomplete, or inaccurate data.

B.4 Determining organizational placement and design

Identifying the value is one thing. Making it happen is quite another. Establishing a successful Competency Center depends on the right planning. Organizations

that take a measured, well-managed approach synergizing people, process and technology are more likely to succeed. Those that do will gain wider support, contribute significant cost savings to the corporate balance sheet and help take business intelligence to the next strategic level.

The Competency Center model varies depending on the need of the organization and its level of maturity, as shown in Figure B-7. It might be an IT-only based initiatives, designed to focus on consolidating the system knowledge necessary to ensure a consistent enterprise strategy for Business Intelligence. It might be Line-of-Business based, to focus on functional business skills and capabilities sponsored by Business Executives, to work closely with IT to bridge the gaps that exist in the organization today. Some Competency Centers are centralized at a corporate head office level, while others are loose networks or regional and divisional teams made up of Business and IT personnel.
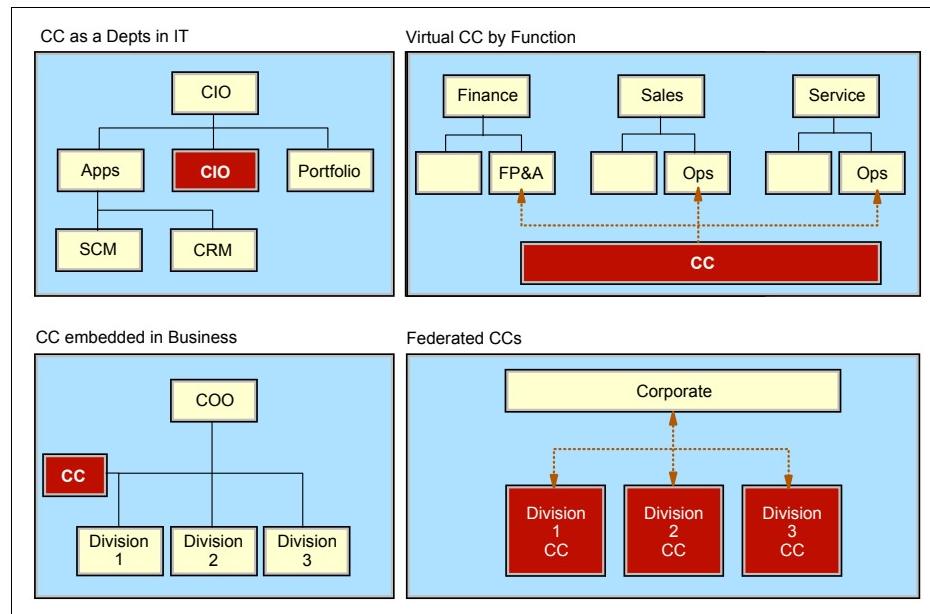
Figure B-7   Competency Center models

The Competency Center design, whether centralized, de-centralized, or based on full time employees or a virtual set of community skills, therefore depends on the functions that it is responsible for and the problems that it is looking to resolve. The right technologies enable a complete, relevant and consistent view of information anytime, anywhere, the right metrics help drive the corporate strategy for growth and profitability, the right organizational design and business model helps the company embrace a shared view of how the company creates,

manages and leverages information to the fullest potential. The common functions of a Competency Center can therefore vary and might include all, or portions of the following functions:

► Best Practices and Standards Management: Providing a clear process and repository for approving and sharing best practices and standards for the enterprise

► Advise and Consulting: Providing a functional area of business with advice, guidance, mentoring and internal consulting so that project teams can become self-sufficient

► Community Services: Where the Competency Center designs and builds content, such as common reports and data packages for use by the broader business communities

► Communications and Evangelism: How the Competency Center communicates and promotes the status, progress, accomplishments, successes, and the overall roadmap to the business at large to help embrace cultural change

► Enterprise Technical Architecture: How the Competency Center directs, builds, and supports the technical infrastructure that supports the needs of the business. It is here that the Cloud Innovations are having most impact in organizations today, whereby the Competency Centers capabilities are amplified by an organizations ability to take advantage of virtualized resources, automated and standardized services, location independence, and rapid scalability / elasticity to meet the ever changing business needs to use technology

► Support: How the Competency Center provides an Information based help desk function to the business, or otherwise supports a broader help desk function

► Education: How the Competency Center trains and educates the business or related IT functions and the best use of the technologies it supports

► IT Governance and Alignment: How the Competency Center plugs into the broader IT Governance processes and Steering Committees, such as those that manage Project and Change Management, Portfolio Management, Vendor Management, and License Management amongst others

► Data Governance: Where the Competency Center creates a suitable framework, or interfaces with existing Data Governance disciplines and programs across the organization and lines of business

► Business Strategy Alignment: How the Competency Center plugs into the Corporate Business Strategy top ensure that the technology-related initiatives are meeting the most important needs and priorities of the business for not only the current initiatives but the longer term strategic goals of the organization and its growth

Consider the following steps to make a Competency Center a reality within your organization:

1. Acknowledge the need, urgency, and willingness for change

   A Competency Centers' design is influenced by a range of factors, including business culture, historical deployment activity, specific business pains and measures of success. As broad as they might seem, it makes more sense to start with a limited focus.

   Consideration for forming a Competency Center that is driven by a need and urgency for change, specifically around the ability to meet business demand for better decisions using Business Intelligence and Analytics technology. Who is feeling the pain? Who is raising the need for change? How is this being expressed? Is there a willingness to institute change, at least in some parts of the organization?

   To make change happen and to make it persistent requires sponsorship at the highest levels. It is vital to have a C-level or VP-level sponsor of the Competency Center initiative because the primary goal is to optimize business performance driven by business strategy and priorities. Without executive commitment to an overall Business Intelligence-related vision, it will be difficult to get buy-in from other stakeholders to utilize information as part of the day-to-day operation and management of the business and to ensure alignment to reach strategic goals. The sponsor must be a firm believer in driving the culture around measuring the business to manage it effectively, to use Business Intelligence as part of the management philosophy, and to ensure that all levels of management are aligned with the vision.

   Executive sponsors should therefore be committed to making change happen. They should be the guiding force to nurture the development, growth, funding and navigation of company culture to drive the Competency Center from idea stage to reality. They will be a primary drivers for building an initial guiding team and establishing critical early momentum.

   Because a Competency Center evolves over time, so, too, must its sponsorship. Sponsors might change depending on maturity, scope and priorities. Once competency is established, the scope can expand. Over time, the Competency Center can play an increasingly strategic and proactive role in the organization. Bear in mind, however, that there is no one-size-fits-all approach to Competency Centers. It all depends on each organization's specific needs and the dynamic considerations of human capital, managed process, culture and technology infrastructure.

   For example, in the early start-up phase, the Competency Centers might be sponsored by the CIO as a catalyst for change. The CIO might see a need for better business partnership, but recognizes that the initial foundation-building for the Competency Center might need to start within IT. This insular

approach helps to quickly improve the data foundation and generate some early wins for the business.

As time progresses, and with foundational elements in place, the sponsorship focus might shift to, for example, the CFO. This extended perspective allows the Competency Center to deepen its investments in broader competencies, including performance management and the integration of historical views – from the data warehouse – with future views to build stronger financial planning, budgeting, and forecasting capabilities.

2. Interlock business and IT with focus on building a guiding team

A Competency Center becomes most critical when there is a recognized need that crosses functional boundaries. To succeed, this cannot be a strictly IT-driven initiative. The business needs to be actively involved in the creation and operation of the Competency Center. In general, if IT tries to push IM into the business, the investments are not necessarily in line with the business needs and the benefits are not always realized or recognized by the functional areas.

This thinking extends deeper into the functional areas of the business. Who are the business users who work with data on a regular basis? Who are the analysts and power users in the business who are responsible for providing content to others or doing deep analytics in support of management? Finally, who are the IT teams involved in supporting the delivery of information?

Whatever answers are returned, recognize that the need will not be the same in every area of the business at any point in time. Those who are feeling the pain now and see the urgency for change are typically prime candidates to be early participants in the formation of the Competency Center. They will help form the initial community and become the guiding team to drive the first steps to success.

The business must drive the investment decisions and priorities. To be successful, the business must partner with IT to ensure solutions, processes and infrastructure align with the organization's strategic investment decisions. With the establishment of effective Competency Center governance processes, with the right representation from the business, the elements are in place to ensure effective communication between the various functional areas and IT. However, IT resources are often constrained, so the governance process should allow for flexibility to meet specific functional requirements by integrating access to non-IT budgets and resources into the process. This maximizes the return on efficient use of IT resources while ensuring that each functional area has sufficient means to meet its needs.

3. Develop a Strategic plan through Organizational Readiness

Competency Centers can take on many forms depending on the needs, size and objectives of your organization. The setup and alignment of people will be

influenced by many factors and becomes easier and clearer as you take the initial steps forward.

An important first step is to review your organization's perception of the Competency Center and the maturity level of the effort (Figure B-8). Keep in mind that some of the pieces might already be in progress, even if they aren't formally identified as being part of a Competency Center. The following questions can help you determine where you can be along the continuum:

– Is a Competency Center a new concept to your organization?

– Do you already have a Competency Center in place and want to improve it?

– Are you ready to take the next step, toward a Competency Center?

– Based on the answers to these questions, you should be able to identify the maturity level of your Competency Center. Moving through these three phases will help you drive toward enterprise success:

– Are you pushing awareness of the Competency Center concept and philosophy through your organization?

– Have you moved on to formalizing your Competency Center plan and service case?

– Are you ready to, or already started the process of, implementing your Competency Center?
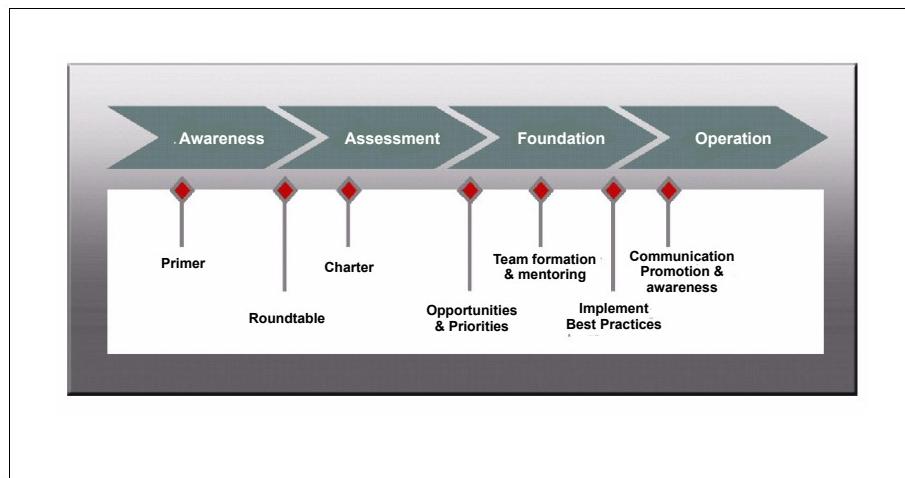
Figure B-8   Determining your organization's perception of the Competency Center and the maturity level of the effort

After you establish the need, have a committed set of executive-level sponsors, and have identified the initial set of business stakeholders, you have the core elements in place to build a Competency Center plan and prepare the organization accordingly. This involves a number of steps, which we illustrate in Figure B-9 on page 329, including:

– Awareness: Ensuring all appropriate groups, including IT and business organizations with a vested interest, are aware of the purpose and value of a Competency Center and the intention to build one within the organization.

– Assessment: With a series of in-depth roundtable sessions and functional maturity assessments, determine the current "as-is" state, the desired "to-be" state, the IM strategy and prioritized roadmap to achieve the desired state within an appropriate time frame, and sanction the Competency Center governance with an approved charter of authority, scope, and roles and responsibilities. This step can require a business case to justify and fund the development.

– Foundation: Start the process to build up and launch the Competency Center including the execution of the initial responsibilities and functions (Best Practices and standards, education), typically in conjunction with a set of information management initiatives/projects. Start communicating to a wider audience within the organization about the Competency Center strategy and roadmap, and that the capabilities and functions are now open for business.

– Operation: Now that the Competency Center is off and running, it is time to consider expansion (more functions, more geographies), improvements or refinements, and opportunities for greater efficiencies (outsourcing some functions). This will also have an impact on where the Competency Center resides, how it is structured, staffed and funded.

Figure B-9 on page 329 shows the necessary steps to build a Competency Center.
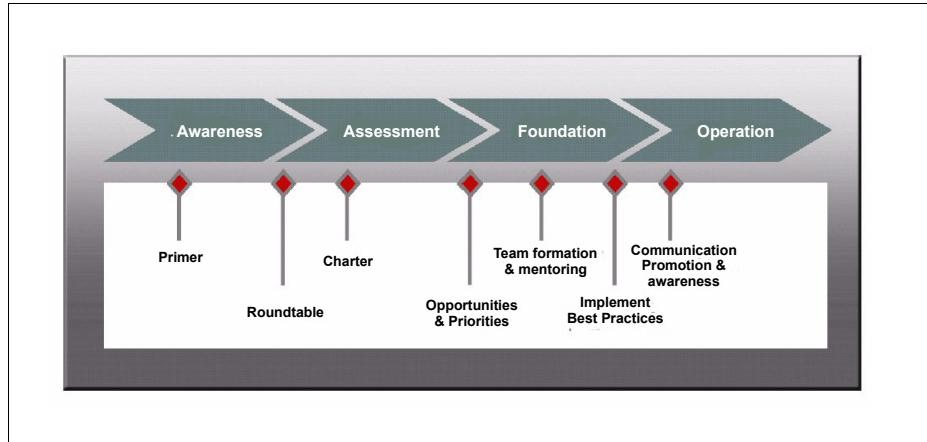
Figure B-9   Steps needed to build a Competency Center

The right selection of people and appropriate skills are critical to the success of your Competency Center– but only after you define the scope. The most successful Competency Centers maintain a mix of technical and business skills, either internal to the consolidated teams or accessible within the organization as required. It is up to the Competency Center to determine which skills, roles and staffing requirements are needed. As these efforts matures and the scope of its function diversifies, additions and changes to the roles and skills will be required to meet the service and capacity needs of the organization.

At an absolute minimum, the Competency Center consists of the following basic roles:

– Competency Center Director/Manager
– Business Analysts
– Technical Consultants

It is important that you clearly define the responsibilities of these roles, including the functions of the Steering Committee, leadership qualities for the Competency Center Director/Manager, and core roles and traditional responsibilities of the team personnel. Staffing the Competency Center should not be viewed as a point-in-time effort. Rather, it is an ongoing competency that will require additional investments in education and certification.

4. Think strategically, but act tactically in incremental steps that add immediate value.

Look for opportunities to bring quick but valuable *wins* to the organization. Projects with long development cycles and extensive scope can impact a

sense of momentum and lower commitment to the Information Management vision.

The most successful Competency Centers are formed through a pragmatic development effort that matures over time. Based on our experience with customers implementing Business Intelligence solutions, consider the following:

– Start small
– Think strategically
– Systematically accelerate

Because Business Intelligence solutions have a cultural impact on an organization, small simple steps will yield greater return as users adopt the Business Intelligence standard, and the Competency Centers success and the projects it services drive subsequent successes.

5. Measure success and recognize value of improvement (Business Effectiveness).

With some early successes accomplished, it is important to then continually measure the benefits that the Competency Center has brought to the organization, both directly in terms of efficiencies and productivity in IT and the business (business process improvements) and indirectly by improving the overall decision making capabilities.

After wins are established and success can be quantitatively measured (more users, improved data quality, more cross-functional views of business performance), those successes should be regularly communicated to the broader community of users, executive management and the business in general. Success will generate interest for new initiatives, further expanding the benefits of Information Management.

B.5 Summary: Winning Conditions

For many who have made investments in technologies like Business Intelligence, Financial Performance & Strategy Management and Analytic / Predictive Applications, deployments have tended to be tactical or localized at the department level. Our studies show that organizations that take a well managed, pragmatic approach to technology are able to unlock the value of business information for competitive advantage.

Although the level of formality might vary from organization to organization, the goal of business models today are designed to:

► Promote and provide delivery enablement through a consistent set of BI skills, standards and best practices

- ► Enable repeatable, successful BI deployments through the development and focus of people, process, and technology
- ► Measure success in terms that are relevant to the entire organization, and not just a single team, department or project

If technology is to extend beyond tactical deployments to become a broader-based solution, a managed, predictable approach is needed. A Competency Center defines the knowledge, standards, and resources needed to make this happen and is essential to the strategic deployment of technology because it:

- ► Maximizes the efficiency, use and quality of your technology across all lines of business
- ► Leads to technology deployments that have higher success and deliver more value, at less cost, in less time
- ► Drives end user adoption to ensure its success (simply providing technology to an increasing number of information consumers doesn't guarantee more people will use it)
- ► Eliminates the gap between Business and IT
- ► Enables business agility and improved technology management which will help to drive business efficiency

Through central, internal network accessed, virtual, web-based environments (for example, Private Cloud), Competency Centers accelerate the large scale user adoption of decision making capabilities, enable enterprise efficiency and deliver real value to the business. Not only implementing powerful web based Business Intelligence and Performance Management solutions, organizations are looking to also take advantage of the benefits of location-independent resource pooling to address challenges in scarce skills and capabilities. Often, organizations are unaware that they have already started taking their first steps into Cloud Computing, namely using Private Clouds.

A Private Cloud managed by a Competency Center provides on demand self-service to enable business users, ubiquitous network access, rapid elasticity for agility, location independent resource pooling & best practices community sharing. Cloud based infrastructure capabilities provide and support self-service, automation, elasticity and variable costing models. This provides Competency Centers with an ability to enables its users to consume IT as a service, without burdening users with intricacies of how the services work behind the scenes.

With a Cloud Environment, Competency Centers recognize higher utilization, economy of scales benefits, lower capital expenses and lower operating expenses, resulting in standardized use of technology including easier access, flexible pricing, reuse and share and ease of integration. This results in an

organizations ability to increase project cycle times, lower support costs, optimize utilization, improve compliance and security and ultimately increases the end-user experience.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

For information about ordering these publications, see "How to get Redbooks" on page 334. Note that some of the documents referenced here might be available in softcopy only.

- ► *Introduction to the New Mainframe: z/VM Basics*, SG24-7316
- ► *Leveraging IBM Cognos 8 BI for Linux on IBM System z*, SG24-7812
- ► *Enterprise Data Warehousing with DB2 9 for z/OS*, SG24-7637
- ► *z/VM and Linux Operations for z/OS System Programmers*, SG24-7603
- ► Achieving High Availability on Linux for System z with Linux-HA Release 2, SG24-7711
- ► IBM System z Strengths and Values, SG24-7333
- ► *Security on z/VM*, SG24-7471
- ► *Security for Linux on System z*, SG24-7728

Online resources

These Web sites are also relevant as further information sources:

- ► IBM Cognos 8 Online Documentation

  http://publib.boulder.ibm.com/infocenter/c8bi/v8r4m0/index.jsp

- ► IBM Cognos 8 Documentation (PDF)

  http://www-01.ibm.com/support/docview.wss?uid=swg27015681&rs=3442

- ► IBM Cognos 8 Business Intelligence Product Overview

  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP10143 7

How to get Redbooks

You can search for, view, or download Redbooks, Redpapers, Technotes, draft publications and Additional materials, as well as order hardcopy Redbooks publications, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

IBM

Redbooks

IBM Smart Analytics Cloud

IBM Smart Analytics Cloud

Understand the business value of a Smart Analytics Cloud

Understand the Smart Analytics Cloud architecture

Implement and manage your own Smart Analytics Cloud

This IBM Redbooks publication presents a Smart Analytics Cloud. The IBM Smart Analytics Cloud is an IBM offering to enable delivery of business intelligence and analytics at the customer location in a private cloud deployment. The offering leverages a combination of IBM hardware, software and services to offer customers a complete solution that is enabled at their site. In this publication, we provide the background and product information for decision-makers to proceed with a cloud solution.

The content ranges from an introduction to cloud computing to details about our lab implementation. The core of the book discusses the business value, architecture, and functionality of a Smart Analytics Cloud. To provide deeper perspective, documentation is also provided about implementation of one specific Smart Analytics Cloud solution that we created in our lab environment. Additionally, we also describe the IBM Smart Analytics Cloud service offering that can help you create your own Smart Analytics cloud solution that is tailored to your business needs.